A Musical Wormhole

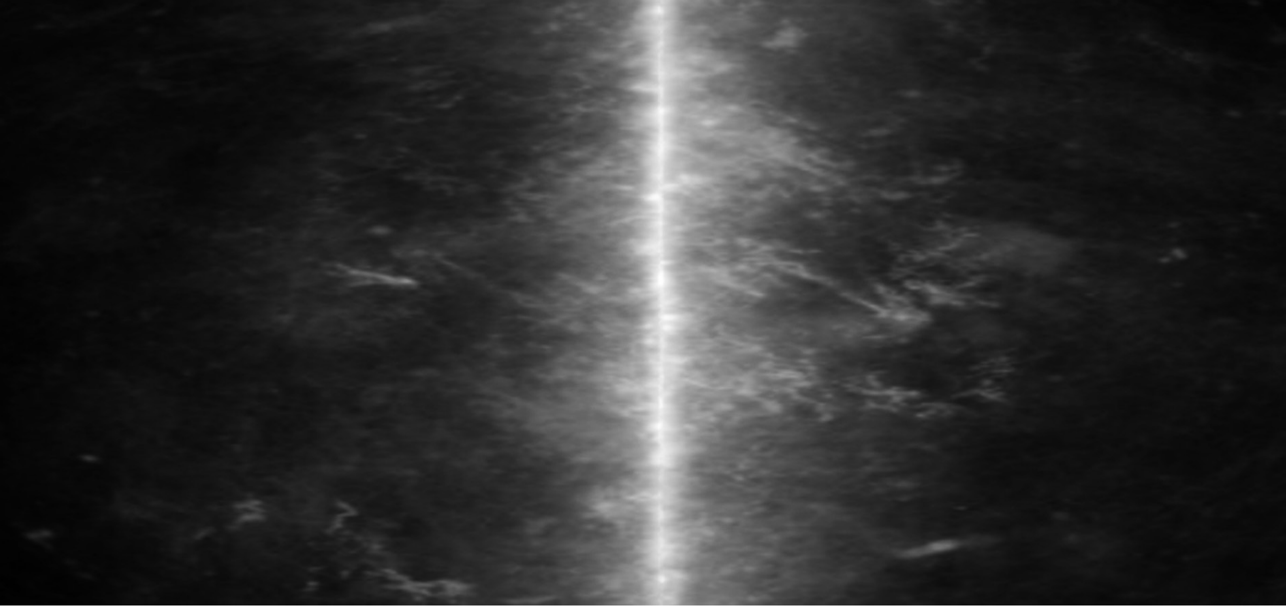
Between the Infinity of the Universe and the Limitations of the Computer Game Media - A Theoretical Basis for Adaptive Vertically Dominated Music in Computer Games

Kristensen, Kristian Sønderlund Ross

DOI (link to publication from Publisher):
10.5278/vbn.phd.hum.00056

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record



A MUSICAL WORMHOLE

BETWEEN THE INFINITE UNIVERSE AND THE LIMITATIONS OF THE COMPUTER GAME MEDIA

A THEORETICAL BASIS FOR ADAPTIVE VERTICALLY DOMINATED MUSIC IN COMPUTER GAMES

BY
KRISTIAN S. ROSS KRISTENSEN

DISSERTATION SUBMITTED 2016

AALBORG UNIVERSITY
DENMARK

A MUSICAL WORMHOLE

BETWEEN THE INFINITE UNIVERSE AND THE LIMITATIONS OF THE COMPUTER GAME MEDIA

A THEORETICAL BASIS FOR ADAPTIVE VERTICALLY DOMINATED MUSIC IN COMPUTER GAMES

by

Kristian S. Ross Kristensen

AALBORG UNIVERSITY
DENMARK

Dissertation submitted

.

Dissertation submitted: October 5th 2016

PhD supervisor: Associate Prof. Martin Knakkergaard,
Aalborg University

Assistant PhD supervisor: Prof. Mark Grimshaw,
Aalborg University

PhD committee: Professor Nicolai Jørgensgaard Graakjær,
Aalborg University

Associate Professor Karen Collins,
University of Waterloo

Professor Ola Stockfelt,
University of Gothenburg

PhD Series: Faculty of Humanities, Aalborg University

ISSN (online): 2246-123X
ISBN (online): 978-87-7112-821-5

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2016

CV

Kristian S. Ross Kristensen has a Master in Musicology from AAU. The Master is focused on adaptive music composition and sound design for interactive media on a practical and a theoretical level and includes courses from the University of Aalborg (AAU), the University of Århus (AU), Danish National Academy of Digital Interactive Entertainment (DADIU) and Iceland Academy of the Arts (LHI). He enrolled in the Doctoral School of the Humanities at AAU in 2013.

Kristian is specialized in the fields of adaptive sound design and music for interactive media, film sound and music, music theory, audiovisual theory and music production. He has worked professionally as a sound artist, musician, composer, film sound engineer and sound designer in Denmark, Germany and Iceland. He has taught master classes in sound design and audiovisual theory at Weißensee Kunsthochschule Berlin and Iceland Academy of the Arts, as well as functioned as a music consultant for the game company CCP, and as a sound consultant and sound designer for the interactive design company, Gagarin.

The past few years, Kristian has, besides being a Ph.D.-student, organized art workshops on subjects such as radio storytelling and hydrophone building as well as produced a series of radio shows on Icelandic national radio about, amongst other things, vertically dominated music.

ENGLISH SUMMARY

The Ph.D. research investigates the implications of implementing adaptive VDM (Vertically Dominated Music) in computer games. VDM is proposed as a term for music that is perceptually dominated by vertically oriented (simultaneously sounding) musical qualities like harmony, timbre and micro-texture, and encompasses musical styles such as sound-mass music, textural music and spectral music.

The thesis presents a structural, aesthetic and functional theoretical foundation that allows for the development of an automated system for the implementation of adaptive VDM in computer games.

A phenomenological, semiotic and music-perceptual framework of understanding, is developed, through which the narrative, semiotic, multi-modal and perceptual implications on the gaming experience of introducing VDM in a computer game are assessed. Here the focus is on VDM's association with narrative phenomena such as outer space, infinity, mystery, divinity and paranormal activity by making sensory accessible the presence of "something" that exceeds the boundaries of the senses.

An initial set of vertically dominated compositional principles is provided based on the findings of the research, and practical experiments are presented as a means for demonstrating the applicability of the developed concepts.

The problem of *verticality* is explored and *auditory verticality* is defined on both philosophical, perceptual and structural levels. An insight into VDM structure identifies a set of general musical characteristics of VDM that may be applied to an automated system. A model for understanding the perception of VDM is developed, which suggests regarding musical listening as a phenomenon of *empathetic co-agency* and provides a scale on which music listening experiences can be positioned in a field between horizontal and vertical dominance. VDM's narrative potential is uncovered in terms of the music's semiotic and perceptual attributes as well as in relation to the general challenges posed on computer game music. Structural and aesthetic implications of computer-generated music are examined and solutions to some aesthetic challenges associated with producing music entirely on the computer are proposed. Possible strategies for horizontal development of VDM are examined including deterministic, indeterministic and stochastic methods, and solutions to the problem of algorithmic musical representation in a VDM context are suggested. Finally, a set of appendices document practical experiments and prototypes that demonstrate the application of some of the covered theory in practice.

The thesis argues that a specialized system for adaptive VDM demands a fundamentally new approach to computer-generated adaptive game music, which draws on not only the unique structural characteristics of VDM, but where also

semiotic, perceptual and narrative issues form the basis for the automated composition of adaptive music.

DANSK RESUME

Denne Ph.d. afhandling undersøger implikationer for implementeringen af adaptiv VDM (Vertikalt Domineret Music) i computerspil. VDM foreslås som en term for musik, som er perceptuelt domineret af vertikalt orienterede (samtidigt lydende) musikalske karakteristika såsom harmoni, timbre og mikrotekstur, og som omfatter musikalske stilarter såsom lydmassemusik, teksturmusik og spektralmusik.

Afhandlingen præsenterer et strukturelt, æstetisk og funktionelt teoretisk fundament, som tillader udviklingen af et automatiseret system til implementeringen af VDM i computerspil.

En fænomenologisk, semiotisk og musikperceptorisk forståelsesramme udvikles, igennem hvilken narrative, semiotiske, multi-modale og perceptuelle implikationer for spiloplevelsen ved at introducere adaptiv VDM i et spil vurderes. Fokus er her rettet på VDMs tilknytning til narrative fænomener såsom det ydre rum, uendelighed, mystik, guddommelighed, og paranormale hændelser samt tilstedeværelsen af ”noget”, der overskrider sansernes begrænsning.

Et forløbigt sæt af vertikalt dominerede kompositoriske principper tilvejebringes på grundlag af forskningsresultaterne, og praktiske eksperimenter præsenteres som et middel til at demonstrere anvendeligheden af den udviklede teori.

Vertikalitetsproblemet udforskes og *auditiv vertikalitet* defineres på både filosofiske, perceptuelle og strukturelle niveauer. Et indblik i VDMs musikalske struktur identificerer en række generelle musikalske egenskaber for VDM, der kan anvendes i et automatiseret system.

En model til forståelsen af perceptionen af VDM udvikles, som foreslår at betragte musikalsk lytning som *empatisk co-agentur* og som præsenterer en skala, hvorpå den musikalske lytteoplevelse kan positioneres i et felt mellem horisontal og vertikal dominans. VDMs narrative potentiale undersøges ud fra dets semiotiske og perceptuelle attributter, og i relation til de generelle udfordringer, der stilles computerspilmusik. Computer-genereret musik’s strukturelle og æstetiske implikationer belyses, og løsninger til signifikante æstetiske problemstillinger foreslås. Mulige strategier for horisontal udvikling af VDM undersøges, inklusiv deterministiske, indeterministiske og stokastiske metoder, og løsninger i forbindelse med algoritmisk komposition’s repræsentation foreslås. Endelig dokumenteres i et sæt bilag praktiske eksperimenter og prototyper, som demonstrerer applikationen af nogle af afhandlingens teorier og koncepter.

Afhandlingen argumenterer for, at et specialiseret system til adaptiv VDM kræver en fundamentalt ny tilgang til computergenereret, adaptiv computerspilmusik, som ikke kun forholder sig til VDMs unikke strukturelle karakteristika, men hvor også semiotiske, perceptuelle og narrative forhold danner grundlag for den automatiserede komposition af adaptiv musik.

ACKNOWLEDGEMENTS

I would like to express my gratitude to everyone who has contributed in one way or another to the realization of my Ph.D. thesis.

A sincere thank you to my two supervisors Martin Knakkergaard and Mark Grimshaw for encouragement, editorial support and committed discussions.

Thank you to staff and students at Aalborg University and the Iceland Academy of the Arts. Much appreciation also to Baldur Baldurson and Jón Hallur from the audio department at CCP and Kjartan Ólafsson for collaboration, inspiration and practical perspectives. Mads Walther-Hansen, Justin Christensen, Hans-Peter Gasselseder, Morten Søndergård, Palle Dahlstedt and all members of the Sound and Music Knowledge Group (MaSK).

I would like to thank my family - not least my parents and their spouses for support. A special thank you goes to my mom for proofreading and much help and insight towards the end of the writing process.

I also owe my sincere thanks to Baldur Snorrason for making the illustrations; Sandra and Lasse Burri Gram-Hansen for help with document formatting; Marteinn Sindri Jónsson and Ragnheidur Harpa Leifsdóttir for collaborations on Icelandic radio, Polar Festival and Lunga Festival; Margrethe and Morten, Nils Wiberg, Christian, Kolbrun, Kristian and Gabi for friendship, discussions and support; and to Hugrún for shining a light at the end of the tunnel.

Dedikeret til min farmor.

TABLE OF CONTENTS

TABLE OF FIGURES

CHAPTER 1. INTRODUCTION

The use of sound-mass music by Stanley Kubrick in some of his films, such as Ligeti's *"Atmosphêres"* (1961) in the science fiction classic, *"2001: A Space Odyssey"* (1968) and Penderecki's *"De Natura Sonoris"* (1966) in *"The Shining"* (1980) sparked a fascination in me some years ago. It seemed to me that this music had a peculiar ability to convey, through music, the presence and qualities of some aspects of Kubrick's films that would otherwise be impossible to portray with the same emotional depth - such as ghosts, clairvoyance, infinity and the fundamental existential mystery of the existence of our Universe. On a personal level, a key motivating factor for the writing of this thesis has been to seek an answer to the question of how some musical styles seem to work so well for this purpose: What is it that seems to link the listening experience to these incomprehensible, ungraspable aspects of the narrative of not only these two films, but of a wide range of both films and computer games in the science fiction, horror and mystery genres?

While doing my masters thesis on computer game music I read a book chapter in which Jesper Kaae writes the following:

> "All music can be seen as both horizontal and vertical, but some music seems to be either vertically or horizontally orientated." (Kaae in, M. K. Collins, 2013, pp. 87)

The vertical and horizontal dimensions refer to the traditional written musical score in which the vertical dimension represents simultaneously sounding pitches while the horizontal dimension represents temporal development. Kaae briefly argues that using non-linear composition techniques may create music that is less dependent on time and thus may be well suited for the non-linear computer game medium. He points to precisely Ligeti and Penderecki as representatives of such music.

While briefly pointing out a potential of non-linear music with a *vertical orientation* for being well suited for the non-linearity of game media, Jesper Kaae does not provide further research on the topic and his mentioning of music as either *vertically* or *horizontally orientated* is left somewhat rudimentary - although inspiring. Kaae provided a connection between the film music used by Kubrick and the computer game media.

In my proposition of the terms *vertically* and *horizontally dominated music* (VDM and HDM), I owe a debt to Kaae's mentioning of *vertical* and *horizontal orientation*. I have exchanged the word "orientation" with "domination" as the latter suggests a certain "battle" between the two - a music-perceptual "fight for territory" that I have come to realize through the elaboration of these concepts in the first half of this thesis. The notions of VDM and HDM are thoroughly introduced in chapter 2, where I define VDM as:

Music in which the expressivity of auditory verticality dominates over that of auditory horizontality.

The inclusion of the word ”wormhole” in the thesis title is not merely associable to the fact that outer space plays a role in the research as a narrative setting. Black holes and wormholes are suitable metaphors for the music-perceptual attributes of VDM, I will argue, because they represent phenomena in which the dimensions of time and space fuse into a *singularity*. A significant aspect of my argument in the thesis concerns the idea that vertically dominated music is related to a similar concept - an auditory space-time singularity - that represents, in the listening experience, an encounter with the incomprehensibility of *continuum*.

My interest in the subject matter of this research is thus partly motivated by a fascination with the use of sound-mass music in the audiovisual contexts of science fiction, horror and mystery genres; it is partly initiated by the proposition made by Kaae of this music’s potential in non-linear media; and it is sparked by a fascination with *verticality* at a deeper level - the philosophical implications of the occurring simultaneity (the *now)* and the incomprehensible continuum. The thesis can be seen as a theoretical investigation into how the relationship between vertically dominated music and the computer game media may be exploited with the purpose of giving rise to sensations, so to speak, through these media, of that which ”cannot be sensed” - thereby giving a voice to the incomprehensible.

In other words, I ask, what is it in vertically dominated music that might allow it to function as a ”musical wormhole” between the infinity of the universe and the limitations of the computer game media?

1.1. CONTRIBUTION TO THE RESEARCH FIELD

The scientific field of computer game music, although continuously growing, is still of a relatively small size due to the young age of the media and a typical emphasis on other aspects of gaming such as graphics, gameplay, social and economical aspects, to name a few. While several other works could have been mentioned, I shall briefly point to a few larger works in this respect. These include: ”Game Sound - an introduction to the history, theory, and practice of video game music and sound design" (K. Collins, 2008); ”From Pac-Man to pop music” (M. K. Collins, 2013); ”Game Sound Technology and Player Interaction: Concepts and Developments” (Grimshaw, 2010); ”A Composer's Guide to Game Music” (Phillips, 2014); and ”The Oxford Handbook of Interactive Audio” (K. Collins, Kapralos, & Tessler, 2014). A terminology exists around game audio – and to some extent also specifically around game music.

The mentioned literature presents a theoretical basis for understanding vital aspects of existing approaches to music in games and provides good impetus for further research. It provides documentation of important issues to which game music must adapt, not least in terms of the media's non-linearity. However, a common symptom of the theory concerned with music in games is an almost complete disregard of vertically dominated musical traits.

While the research field of computer game music is growing, the focus seems to be exclusively on what I refer to as horizontally dominated music (HDM). Not only is the focus directed at HDM in academia, but, as became evident during my collaboration with the game company, CCP, also the available tools for implementing adaptive sound and music in games seem to have a focus exclusively on solving issues arising from horizontal dominance. Until now, neither academia nor the computer game industry has directed any serious attention to the unique functional, aesthetic and narrative potentials and challenges posed by introducing adaptive VDM in games.

This thesis provides the theoretical foundation and initial practical perspectives to fill this gap, both in academia and in the industry. It does so by exploring philosophical, aesthetic, structural, perceptual, narrative and technological perspectives on the subject. During these explorations it has become clear that existing frameworks for understanding music in games within these perspectives are insufficient to account for VDM - both as a music compositional technique and as a musical percept. This has led me to propose a series of expansions in the thesis of existing theories as well as necessary inventions of new frameworks of understanding that are capable of managing the unique structural and perceptual attributes of this music. Furthermore, the considerations associated with adaptive VDM in games, which I put forth in this thesis, positions VDM in the broader context of music and sound per se. This means that part of the provided theoretical foundation may be applied broadly across other styles of music, as well as relate, in some cases, also to sound and music in linear media.

1.2. PROBLEM STATEMENT

This Ph.D. thesis investigates possibilities for the generation of adaptive music in computer games that have as part of their narrative such phenomena as outer space, infinity, mystery, fear, divinity, paranormal activity - or generally, a presence of "something" that exceeds the boundaries of the human senses. Its purpose is to address practical and theoretical implications of implementing vertically dominated music in computer games in order to find new solutions to the challenges posed by the non-linearity of computer game media.

The investigations of the Ph.D. research can be categorized into the following four principal enquiries:

1. **How may a structural, aesthetic and functional theoretical foundation be devised that allows for the development of an automated system for the implementation of adaptive VDM in computer games?**

2. **What does the construction of an initial set of vertically dominated compositional principles on the basis of such a foundation entail?**
3. **How may a phenomenological, semiotic and music-perceptual framework of understanding be developed, which is capable of explaining the narrative, semiotic and perceptual implications on the gaming experience of introducing VDM in a computer game?**
4. **How can such a framework be deployed as an integral part of the development process of an automated system for implementing VDM in computer games?**

These four principal enquiries can be expressed as three main objectives of the research:

1. **To create a structural, aesthetic and functional theoretical foundation from which an automated compositional system can be built.**
2. **To develop a phenomenological, semiotic and music-perceptual framework of understanding, and through this framework, assess the narrative, semiotic and perceptual implications on the gaming experience of introducing VDM in a computer game. The focus is here on games that have as part of their narrative such phenomena as outer space, infinity, mystery, fear, divinity, paranormal activity - or generally, a presence of "something" that exceeds the boundaries of the human senses.**
3. **To provide an initial set of vertically dominated compositional principles based on the findings of the research as well as as a means for demonstrating the application of the developed concepts.**

1.3. METHODOLOGY

The overall methodological approach is qualitative and founded on the synthesis of existing knowledge as a basis for the invention of new ideas. The method applied in the thesis combines theory from a wide range of fields including: Musicology; music technology; sound and music perception; cognitive psychology; neuroscience; astrophysics; audiovisual theory; multi-modal perception; theory of game sound, game music, film sound and film music; and algorithmic composition.

A range of new theories and ideas are presented out of necessity as research into VDM and its implications in context of computer games has not previously been covered in academia to any serious extent. The theoretical approach, which paves the way for these ideas in the thesis, is characterized by both structural and phenomenological perspectives - and precisely the relationship between musical structure and experienced musical structure plays an essential role throughout the thesis.

Although the research is highly qualitative and theoretical, practical experiments, involving the algorithmic composition program, CALMUS, the space-based MMO, EVE Online and my own development of prototypes for VDM generation, have served as inspiration to many of its ideas and conclusions. These experiments are treated as cases through which my theories have been tested in practice, and based on which, not least, some theories have been inspired. No systematic quantitative collection of data from these experiments has been sought, and no systematic analyses of experiment results have been attempted in a quantitate manner. This being said, the appendices A, B, C, D and E describe the experiments and their role in the research in detail.

Although dealing with a range of idealistic ideas in regards to musical structure, it's perception and philosophy, this thesis should not be read as an attempt to achieve an utopian musical expression. At the heart of the dichotomy of horizontal and vertical dimensions in music lies the paradox that purely vertical sound is impossible. There *is* no sound if time is not passing. I thus use the term *vertically dominated music* rather than, for instance, *vertical music*. In order to explain the workings of VDM, however, it is necessary to present ideas that may seem idealistic in the name of conveying the entire scale of vertically dominated compositional possibilities.

Furthermore, even though VDM does, in most cases, contain a degree of horizontally dominated musical traits (such as melodic fragments, hints of causal direction and occasional musical gestures), both academia and the computer game industry already have the theoretical and practical tools to understand and realize such horizontally dominated musical structures in the context of games. Any horizontally dominated musical structures inside a vertically dominated musical expression needs special attention associated with HDM. It is not the aim of this research to cover this subject in any detail as it is discussed extensively elsewhere.

Rather, my focus in this thesis is put on the implications that are unique to VDM and must pave the way for, especially, the most extreme forms of vertical dominance to be understood and realized musically.

1.4. THESIS STRUCTURE

My investigations are divided into 8 primary chapters (not counting the introductory and concluding chapters).

Chapter 2 serves as a theoretical substrate for many aspects of the thesis. Its contribution is focused mainly on **enquiry 1 of the problem statement.** It is concerned with the problem of *verticality* on a philosophical and perceptual level.

The first half of the chapter deals with the problem of auditory verticality and, in addition to a short overview of some historical developments in verticality in western music tradition, includes the subjects of: Time scales in music; basic units of auditory perception; musical time conceptions; auditory and musical stasis; as well as perceptual windows as an explanation for the experience of stasis. Three new musical time conceptions are proposed in this part of the chapter: *Non-cyclic vertical time*; *cyclic vertical time*; and *irreversibly changing vertical time*; and a new time scale is proposed: the *vertical extension* time scale. The second half of the chapter is concerned with discrete and continuous space-time and covers: Bergson's notion of *pure duration*; Boulez, Deleuze and Guattari's opposition of *smooth* and *striated* space-time; gestalt principles; auditory scene analysis; as well as a look at striation and perceptual windows from the perspective of gestalt theory. Chapter 2 concludes with providing definitions of the notions: *vertical extension*; *vertical expression*; *auditory verticality* and *auditory horizontality;* as well as *vertically* and *horizontally dominated music* (VDM and HDM).

Chapter 3 addresses subjects that relate primarily to **enquiries 1 and 2 of the problem statement**. It deals with the structure of auditory verticality and follows up on the exploration of musical time carried out in chapter 2 by addressing musical space. The chapter divides the structure of auditory verticality into tone systems, harmony, timbre and texture. Subjects covered include: tonal key; harmonic base interval; tone clusters; consonance and dissonance; timbre and color; multi-dimensional timbre-space; and practical simplifications of the notion of timbre. The chapter further looks at musical texture according to the time scales of music as well as Denis Smalley's concepts of spectromorphologies; internal and external note-views; and textural motion. Lastly a brief discussion on acoustic properties and sense of space as an aspect of auditory verticality is presented. Chapter 3 provides the main contribution to a list of vertical parameters that can be found in Appendix F.

Chapter 4 is likewise concerned primarily with **enquiries 1 and 2 of the problem statement**. In this chapter, I turn my attention to the structure of VDM. Common traits of VDM are identified in the form of three primary characteristics that govern the music. An overview of expansions of vertical expressivity is provided; the notion of *musical entropy* is proposed and discussed; and a perspective on VDM structure as *rhizomatic* is suggested. The chapter goes on to present insight into Geörgy Ligeti's *"Atmosphêres"* and *"Lux Aeterna"* by providing a summary of analyses of the parts of these works that were used in Kubrick's *"2001: A Space Odyssey"*. Some compositional techniques of Penderecki and Xenakis are highlighted before taking a look at the methods, concepts and techniques of spectral music.

Chapter 5 is concerned primarily with **enquiry 3 of the problem statement** by addressing the perception of VDM. In this chapter, I suggest a framework for understanding the perception of music based on musical anticipation and properties of the human mirror neuron system. Subjects covered include: Retention and protension; cognitive schemas; and musical past as historic form. I introduce in the framework the concepts of *historically based anticipation, equilibrium-based anticipation* and *perturbation-based anticipation* (or past-listening, future-listening and now-listening respectively). I discuss the mirror neuron system as a shared neurofunctional substrate for motor-action, language, emotion and music. I introduce in this chapter a number of new ideas and concepts which may explain *music listening* as governed by what I refer to as *empathetic co-agency*. The chapter concludes by discussing ”black holes” as a metaphor in music and sound perception as well as introducing the idea of a musical *event horizon*.

Chapter 6 contributes to answering **enquiries 1, 3 and 4 of the problem statement** by exploring the narrative potential of VDM. It includes a discussion of the applicability if the terms, *diegetic, non-diegetic* and *trans-diegetic* with regard to computer games. I introduce the concepts of *overt* and *covert* trans-diegetic music and argue that the non-diegetic narrative layer has special narrative potentials, also in computer games. The chapter further looks at VDM in the context of multi-modal perception. Media competence and VDM's potential for narration through *semiotic attributes* as well as *perceptual attributes* is discussed. I propose the terms *game-external* and *game-internal* competences and relate these to VDM's general and specific relation to its multi-modal context. Concept development for trans-diegetic multi-modal connections between VDM and the game in which it is implemented is covered. Chapter 6 further looks at outer space as a narrative setting. Subjects within this section of the chapter include: VDM and space oriented science fiction; sounds of the Golden Age of space-oriented science fiction; as well as infinite reverb and the problem of sound in space. I briefly discuss in this chapter the characteristics of what I call ”benign” VDM and conclude by arguing that *outer* space, based on the findings of this and the previous chapters, may perhaps best be conveyed through a breaking with the intrinsic boundaries of an *inner* space.

Chapter 7 contributes primarily to **enquiries 1 and 2 of the problem statement** by addressing general implications of sound and music in computer games. It is divided into five main topics: General challenges of game sound; adaptive music; immersion and avoiding competition for player attention; functions of game music; as well as a last section which relates VDM to the challenges of computer game music. The subject matter of this chapter includes: Technological frame, game genres and non-linearity; *direct* and *indirect* adaptive music, common techniques for adaptive music, *arborescent* and *rhizomatic* adaptive music and musical variation in adaptive music; functions of music in the computer game medium, functions of sound in an acoustic ecology, and functions of film music.

Chapter 8 is concerned mostly with **enquiries 1 and 2 of the problem statement** by addressing computer-generated music. The chapter is divided into three primary section covering: Algorithmic composition; challenges of generative music in games; and approaches to adaptive VDM in games. The chapter discusses some terminology of algorithmic composition as well as a range of traditional algorithmic techniques in the categories of transformational and generative algorithms. The notion of *musical coherency* is discussed in a VDM context and a concept of the *complexity of the sounding musical manifestation* is proposed. *Music-shaping stages* as an important aspect of identifying aesthetic shortcomings of traditional algorithmic methods is addressed and a framework for categorizing music algorithms based on their *basic unit of composition* is proposed. The chapter concludes by suggesting that VDM demands additional algorithmic categories to be invented: *Vertical extension-transformational* algorithms and *micro-generative* algorithms. It concludes by assessing that a vertical extension-transformational algorithm with limited micro-generative functionality may be the best suited approach for adaptive VDM in games.

The topic of chapter 9 is horizontal strategies for adaptive VDM. The chapter is concerned with **enquiries 1, 2 and 3 of the problem statement**. It addresses the crucial, and not yet covered, aspect of VDM composition that is concerned with the horizontal development of the music. Here VDM horizontality is discussed in terms of different time scales at which it may take place. The musical representation of VDM is seen as multi-dimensional envelopes that can be either *static* or *adaptive* and used for either *composing* music through vertical parameters or *superimposing* textural micro variations onto musical material that is itself created elsewhere. I argue that horizontal development in computer game music may be either *gameplay-dependent* or *gameplay-independent*. Finally this chapter discusses three different methodologies for horizontal development of VDM: *Deterministic, indeterministic* and *stochastic* methods.

The concluding chapter 10 summarizes the findings of chapters 2-9 in the context of the problem statement.

1.5. INTRODUCTION TO APPENDICES

In addition to the chapters, I include in the thesis a series of appendices that document practical experiments done in the context of my ongoing research collaboration with Kjartan Ólafsson (inventor of the algorithmic composition program CALMUS), and the audio department at the game company CCP (developers of the MMO, EVE Online) as well as in the context of my personal sound art and music production endeavors.

The two prototypes, NowEngine and KPS, are included for demonstration purposes and are, at the time of writing, somewhat unpolished. They have come to build on many of the principles associated with adaptive VDM and I include them here for demonstration rather than as a proof of concept. That being said, a number of observations have been made on the basis of these prototypes, which grant the theoretical elaborations of the thesis a welcome practical perspective.

Appendix A describes the prototype, NowEngine, which is built with the purpose of generating extended vertical expressions. The NowEngine can be found on the DVD accompanying the thesis (Appendix G). Appendix B documents my involvement in the ongoing project of implementing CALMUS in EVE Online. As part of the collaboration, I produced a sample library of irregular samples (titled IrregularOrchestra), which can also be found on the DVD. Appendix C describes four linear VDM pieces that I have composed with an initial intention of implementation in EVE Online. Appendix D covers the prototype, KPS (KeynotePlaybackSystem). KPS demonstrates a range of concepts and ideas developed in the thesis and functions as a transformational algorithm for adaptive VDM on its own right. Appendix E contains a brief documentation of experiments done with implementing VDM in EVE Online through the middleware program, Wwise (Audiokinetic).

Multimedia material for all of the above appendices can be found on the DVD, which is labeled as Appendix G. The sound files and prototypes included on the DVD are meant for demonstration and experimentation purposes only. They are not published along side the thesis.

Appendices A, B, C, D and E can be read sequentially as a presentation of practical experiments with strong relations to the theory of the thesis chapters. This is where the deployment of the framework of understanding put forth in the thesis is most elaborate. The five first appendices could have been written as a single self-contained chapter. However, due to their role as a practical perspective on many of the subjects covered in the thesis, I have chosen to label these experiments as appendices so that they may be easily consulted when referred to in the thesis chapters. The appendices are interconnected in the sense that the development of the material they document has taken place as a cause of events in which one experiment led to the initiation of the next. I used the NowEngine prototype documented in Appendix A to produce the sound library mentioned in appendix B. I in turn used this sound library to compose the four linear VDM pieces described in appendix C. These linear pieces were later cut up into fragments that were fed into the KPS prototype as means for demonstrating transformational algorithm for adaptive VDM. And the KPS prototype was developed based on experiences of certain insufficiencies associated with implementing VDM through Wwise discussed in Appendix E. KPS was built as a response to these inadequacies and features precisely the functionality that would have been useful in the implementation of VDM, but which Wwise did not offer.

Finally, Appendix F contains a list of vertical parameters identified in the thesis, and Appendix H presents a glossary of terms and concepts developed through the research.

1.5.1. A NOTE ON RUNNING THE PROTOTYPES

NB: The prototypes only run on mac computers. Users will need to set the security settings in OSX to accept running applications from "Anywhere", as the default setting allows only programs from the "Mac App Store and identified developers". Otherwise it will not work, and you will get an error message saying that the file is damaged. To change the settings do the following:

Apple Menu ->

System preferences ->

Security and Privacy ->

General tab ->

Allow apps downloaded from "Anywhere"

The procedure is also explained here: http://www.tech-recipes.com/rx/45404/mac-downloaded-app-is-damaged-and-cant-be-opened-error-solved/

While the prototypes are not published along side the thesis, future versions may be. I therefore kindly ask readers not to pass the prototypes on to others without permission

CHAPTER 2. THE PROBLEM OF VERTICALITY

2.1. INTRODUCTION

As an initial step in the effort to formulate a structural, aesthetic and functional foundation on which a generative system for adaptive vertically dominated music can be built, it is necessary to first address the fundamental question of *verticality* per se. The problem of auditory verticality is in many ways essential in regards to the formulation of compositional principles for adaptive VDM in games as well as a framework for understanding VDM's impact on the gaming experience. Issues covered in this chapter will form the initial part of the theoretical basis for the development of a broader theoretical model of the relationship between musical structure and musical anticipatory listening states presented in chapter 5, which deals with the perception of VDM.

A variety of musical style categorizations including those referred to as *sound-mass music, sonorism, textural music, spectral music* as well as some examples of *ambient, noise* and *minimal music* have certain musical characteristics in common that set them apart from other styles. They can be seen as governed by a predominately vertically oriented musical expression. This means that musical structures that are primarily oriented horizontally, over time, such as melody, chord progressions, pulse, rhythmical patterns and anything else that lets the listener predict the further flow of the music are either not present or hidden in the music. This may arguably leave the listener with primarily the sonorous ”now” - the vertical musical dimension - as the object for musical listening attention, and put the listener in a listening state where past and future are of comparably less significance. Music which exhibits this domination of the vertical musical dimension I will refer to as *vertically dominated music* or *VDM*.

It is not without problems, however, to separate the vertical and horizontal dimensions in regard to sound - or in regard to any other matter. After a short historical outline of the evolution of verticality in western music tradition, this chapter deals with the paradoxical question of auditory verticality through two primary sections titled: ”The Problem of Auditory Verticality” and ”Discrete and Continuous Space-Time”.

In the first section I discuss briefly the concept of ”sound” as it is used to describe popular music followed by an investigation of the time scales associated with sound and music. Here I address what is referred to by Daniel Stern as *basic units of perception*, 9 time scales of music proposed by Curtis Roads, musical time conceptions including static and dynamic time as well as 5 types of musical time, suggested by Jonathan D. Kramer. A brief look at the astrophysical concept of *the arrow of time* enables me to identify a necessity to expand Kramer's model by appending a musical time conceptions, I call *irreversibly changing vertical time*. Additionally, Kramer's notion of *vertical time* is differentiated into *cyclic* and *non-cyclic* vertical time in order to be able to better distinguish between different forms of static music. Furthermore, the relationship between these added musical time categories and the notion of *stasis* leads to the proposal of four types of stasis: *Cyclic* and *non-cyclic musical stasis* as well as *cyclic* and *non-cyclic auditory stasis.*

This section of the chapter concludes with a look at *perceptual windows* as an explanatory model for the perception of stasis and discusses what Marc Wittmann refers to as *functional moments, experienced moments* and *mental presence.*

In the second section of the chapter the relationship between discrete and continuous structure are examined. These can be associated with perceptual distinctions between Bergson's *quantified time* and *pure duration* as well as the notions of *striated* and *smooth space-time* originally coined by Pierre Boulez and taken up by Deleuze and Guattari. Additionally, *gestalt principles* initially associated with especially the visual domain provide insight into aspects of audition as demonstrated by Albert Bregman's research into auditory stream segregation. Discussions include the relationship between striation, perceptual windows and the formation of auditory gestalts.

The chapter concludes by defining a series of concepts that will have crucial significance throughout the thesis. I thus propose a definition of *auditory verticality* as a perceptual phenomenon constituted by a *vertical expression* and residing within the temporal window of a *vertical extension.* Furthermore, the findings of the chapter pave the way for a definition of *vertically dominated music* (VDM) as well as *auditory horizontality* and *horizontally dominated music* (HDM).

2.2. HISTORICAL PERSPECTIVE

The history of verticality in western musical tradition can be illustrated by a gradual deepening of pitch-space into spectral space - or the *wide open sound world* as Denis Smalley puts it (Smalley, 1997). That is not to say that compositional decisions based on spectral content have not been present from the beginning. Instrumentation is an example of a spectrally based decision. But the radicality and control by which vertical musical properties are manipulated compositionally as well as the significance given to the vertical dimension per se have increased gradually - in some musical styles more than others.

With polyphony as an expression of multiple simultaneous voices, it follows that vertical structures are formed in pitch-space. Written sources of polyphony in Gregorian chant music date back to around 900 AD (M. G. Knakkergaard, Finn, 2003). Before this period, tradition dictated the use of only specific liturgic melodies, but a gradual expansion of these melodies was eventually tolerated by the church - first horizontally through, for example, additions to their termination and later vertically by adding additional voices. It was ensured that the original Gregorian melodies (vox principalis) were always respected in the process. At the earliest stage, polyphony was based mostly on parallel **4ths** and **5ths** although **2nds** and **3rds** would also occur in some cases at phrase beginnings and endings. This kind of parallel singing likely trickled into the liturgical tradition from folk music and does of course not confine itself to Western culture. Polyphony at its origin seems to stem from intuitive improvisation rather than invention. With Ars Nova and the introduction of bar lines in the written representation of music, a more precise timing of polyphonic synchronicity was made possible giving rise to greater polyphonic expressiveness.

The evolution of verticality in western music history is characterized by a movement from simple polyphonic relations in Gregorian chant music through the development of the functional harmonic system - a system, which gradually evolved into an increasingly liberal chromaticism in the romantic era, and later, with serialism, a disintegration of tonal hierarchy and perceivable causal relations (Bernstein, 1976). This evolution had by the 1960s given rise to a sort of radical verticality that is concerned not only with harmonic relations between pitches, but which allows music to have as its most significant trait the harmony, timbre and texture of the sonorous simultaneity.

2.3. THE PROBLEM OF AUDITORY VERTICALITY

The often-used dichotomy of *horizontal* and *vertical* dimensions in sound and music is paradoxical. Sound is by definition subject to duration or horizontality because it is a process. Sound is a function of time and space and any attempt to isolate one dimension from the other is categorically impossible without the sound ceasing to exist. It is evident, however, that a distinction between the two is possible in human perception of sound and useful in both analytical and compositional contexts. When designing sound we can work with the *spectral balance* - a vertical phenomenon - separately from the *dynamic envelope* of a sound - a horizontal phenomenon. In music, vertical parameters such as harmony, dynamics and timbre oppose horizontal ones like melody and rhythm. Music which is mainly horizontally dominated - as by far most music is - is therefore music in which it is the horizontal developments such as melody, chord progression, musical motives, themes, imitations etc. that are the primary subjects for musical attention. Contrary to this, vertically dominated music refers to music where the harmony, texture, dynamics and timbre at any given moment - the occurring simultaneity - are the most important subject for attention. While horizontally dominated music with its cause-and-effect relations deliberately encourages concrete expectations in the listener, which are then met or violated (or partially met and partially violated), the continuation of the musical flow of vertically dominated music is much more difficult - ideally impossible - to predict. But what is *verticality* in regard to sound? What does it consist of? What is the audible content of auditory or musical verticality? How can it be defined as a phenomenon of its own, were we to try? And what would be the duration in which this content unfolds?

2.3.1. VERTICALITY AND ”SOUND” IN POPULAR MUSIC

In popular music, the notion of *sound* holds some connotations that can be considered vertical. Erik Brolinson and Holger Larsen write:

> ”We consider sound a musical dimension comprising those aspects of the musical structure, which are coupled to the experience of a musical "now." This "now" is not a point on the axis of time; the experience of music categorically presupposes a course of time. The duration of a time segment, which can be termed a musical "now" is dependent upon several factors, primarily the musical context. […] Thus, we will suggest a definition where "sound" is the fundamental **character** of all musical elements as it appears in a very short time segment of the music, but which leaves its mark on a longer coherent section.” (Translated from Swedish, Brolinson & Larsen, 1981, pp. 181).

This definition of *sound* is to be understood specifically within the discourse of popular music. It is defined as being a ”fundamental character of all musical elements” contained within a musical "now", but apart from declaring it to have a ”very short” duration, Brolinson and Larsen do not elaborate further on how to determine the duration of this "now" or how it may be clearly defined. Nor is it clarified what is actually meant by ”fundamental character” of musical elements - whether this refers to spectral properties, tonal language, dynamic properties, chord progressions etc.

The quote deals with a concept of ”sound” that is known from some discourses of popular music in which describing a certain band’s ”sound” may refer to a hybrid concept of both musical style and the sonorous characteristics of music production. Although taking place within a musical ”now”, Brolinson-Larsen’s conception of ”sound” suggests a dependency on the broader musical context - a broader temporal scale - because of the definition’s emphasis on the ”fundamental character” being present in ”a longer coherent section”.

While suitable perhaps in the discourse in which it is conceived and useful to some extent as an explication of the volatile musical phenomenon it seeks to highlight, such dependencies on past and future correlate poorly with the notion of musical *verticality* as conceived in the current chapter. The subject matter of this thesis demands a much more concise investigation into the temporal implications of sonorous simultaneity.

2.3.2. TIME SCALES AND VERTICALITY

2.3.2.1 Basic units of perception

A significant time interval in this regard is 50 milliseconds. This is the duration of one period of a 20 Hz sound wave. It represents the point of transition between two very different auditory phenomena - that of rhythm and that of pitch. In other words, humans have a natural ability to perceive regular intervals longer than 50 ms as a horizontal succession of separate events while intervals shorter than this, the threshold of the human hearing range, are experienced as one unified vertical entity, a pitch. While the timespan of 50 ms is clearly significant, from an auditory perspective this duration seems too short to allow for textures caused by, for instance, violin tremolos, snare rolls, the guttural vibrations of a tiger roar or the snoring of a coffee machine, to fully unfold. All of which could arguably be considered part of a sound's verticality as opposed to its horizontal characteristics like envelope, phrasing or thematic gesture.

A certain consensus seems to exist, which suggests that a duration in the whereabouts of 2-5 seconds has a special importance in regard to the duration of a present moment in areas such as language, vocalization, movement and music. This is, for instance, a typical duration for a human breath cycle, a verbal phrase, conversational turn taking, musical phrases and a number of other grouping scenarios (Stern, 2004). The notion of the present moment is here seen as an expression of basic grouping of smaller structures into higher order units. Daniel Stern interestingly remarks, however, that:

> ”Humans can perceive separate events in a sequence that last only between 20 and 150 milliseconds. These are the basic units of perception. But in themselves they do not make life meaningful. […] If we considered each such perceptual unit as a potentially important and meaningful event requiring attention and awareness, it would be like continually being under the fire of a machine gun. These sequences must get chunked into larger units more suited to adaptation.” (Stern, 2004, pp. 41)

The grouping of impressions into larger ”meaningful” gestalts is seen to happen on the basis of shorter event structures - basic units of perception, which represent a potential informational overload unless organized into higher level units. This process could be regarded an expression of verticality as it unifies separate horizontal events into a ”oneness” - an example of which is the phenomenon of pitch arising from the unification of separate sound pulses.

2.3.2.2 Time Scales of Music

Curtis Roads outlines nine time scales of music; *Infinite, supra, macro, meso, sound object, micro, sample, sub sample* and *infinitesimal* (Roads, 2004). Of most relevance to the context of auditory verticality are the *meso, sound object* and *micro* time scales. The present moment as described by Stern would fit under Roads' *meso* time scale, which is characterized by groupings of sound objects into phrases, sequences, melody, counterpoint, themes etc. In other words, Stern's present moment and Roads' meso time scale are capable of containing organizations of smaller perceivable and separable structures, such as notes and motives, inhalation and exhalation, spoken sentences and so on. The *basic units of perception* on the other hand, occurring in the area between 20-150 ms, which are suggested by Stern to perceptually group together out of a certain cognitive necessity, fall within the *micro* time scale. This covers: *"Sound particles on a time scale that extends down to the threshold of auditory perception"* (Roads, 2004, pp. 4) embracing transient audio phenomena, the transition between rhythm and pitch, guttural tiger roar vibrations, snare rolls etc. and extends all the way down to the threshold of timbre perception at about several hundred microseconds. This leaves a gap between 150 ms and 2 seconds, which roughly corresponds to Roads' *sound object* time scale. Its definition can be summed up as:

> "A basic unit of musical structure, generalizing the traditional concept of note to include complex and mutating sound events on a time scale ranging from a fraction of a second to several seconds." (Roads, 2004, pp. 3)

It is within this time frame that phenomena having to do with, for example, the deliverance of a musical note occur such as vibrato, tremolo, timbre variations and dynamic and spectral envelope of a tone's onset, continuation and termination. The term *sound object* time scale is potentially problematic in relation to musical verticality as it takes sides in the question of whether or not the involved structure can be independently segregated from the whole - a notion that will be of critical importance in my argumentation of this chapter.

It should be mentioned here that the often used but vaguely defined notion of *musical texture* may refer to structure on any of the meso, sound object and micro time scales and have quite different perceptual and structural implications for each. Musical texture will be taken up in chapter 3.

2.3.2.3 Musical Time Conceptions

Two important philosophic perspectives on time are that of *static time* and *dynamic time.* Put very shortly, the static conception of time can be illustrated by the image of the placement of events on a time line where they act as points situated simultaneously, before or after each other. In the dynamic time view, time is symbolized by different states in the form of past, present and future. The static time line seems to be observed from somewhere "outside" the now, while the dynamic time is observed from "inside" (M. Knakkergaard, 2007). The experience of music is seemingly capable of drawing on both time perceptions. Martin Knakkergaard writes:

> ”…where a tension/release principal is present, whether in regard to harmony, rhythm, meter, dynamics, texture, articulation or the like, dynamic time governs, while a static conception of time is present in musical styles with an emphasis on the monotone, repetitive and seemingly tensionless and static.” (Translated from Danish, M. Knakkergaard, 2007, pp. 148)

According to Knakkergaard’s remark, concrete experienced musical structures such as those giving rise to the presence or absence of causality can be seen as belonging to an understanding of time which is dynamic or static respectively. In other words, dynamic time can be coupled to directed *process* towards a goal, while static time is expressed by a lack of such direction.

Goal orientation is essential in Jonathan D. Kramer’s formulation of 5 types of musical time: *Goal-directed linear time, non-directed linear time, multi-directed linear time, moment time* and *vertical time* (Kramer, 1988, ch. 2). I will not go into these 5 types of musical time in great detail, but merely point to a few useful considerations in Kramer’s categorization. Moreover, Kramer’s formulations of linear and non-linear musical time shed light on the potential usefulness of the dichotomies static/dynamic time as well as linear/non-linear time in regard to a music-perceptual model, which will be further elaborated upon in chapter 5 of this thesis. Let us first take a closer look at the concept of linearity.

> ”Let us identify linearity as the determination of some characteristic(s) of music in accordance with implications that arise from earlier events of the piece. Thus linearity is processive. Nonlinearity on the other hand is non-processive. It is the determination of some characteristic(s) of music in accordance with implications that arise from principals or tendencies governing the entire piece or section.” (Kramer, 1988, pp. 20).

Musical linearity is here presupposed by the music referring back to itself, so to speak, by causal direction. In this quotation, Kramer addresses musical linearity through a structural approach and does not take into account the question of whether or not this structure is actually registered by the listener. Linear musical time is parallel to the dynamic time view described by Knakkergaard, while non-linear musical time associates itself with a static understanding of time. The linear musical time is further differentiated into three types according to the degree of experienced goal orientation: *Goal- directed linear time, non-directed linear time* and *multi-directed linear time.* As the linearity of these three musical temporalities defines them as having causal direction, the mentioned discrepancy between compositional structure and *perceived* structure is precisely identified by the addition of *goal-directed, non-directed* and *multi-directed.* The Western musical tradition is to a large extent founded on a sense of *goal-directed linear time*, exemplified especially by the tonic as a harmonic goal, but other musical goals are relevant as well, such as the resolution of dissonance to consonance, musical gesture and other forms where musical tension is resolved to an equilibrium. In Western music culture, *non-directed linear time* can be exemplified by atonal music. In atonal music there is linearity because compositional structures refer back to themselves (e.g. by inversion and retrograde of a tone row), but the structure is not directed towards a predictable goal. In *multi-directed linear time* the musical structure is aiming at several goals at once, which makes it difficult for the listener to determine which goal is being pursued.

Kramer differentiates *non-linear* time into *moment time* and *vertical time*. For both of these, their non-linear nature dictates an absence of causal direction. However, while the non-linearity of moment time especially exists as an indifference towards the sequential order of musical sections or *moments,* vertical musical time is characterized by being devoid of experienced causality all together - and thus by a very strong preference for the vertical dimension rather than the horizontal. Inspired by Kramer's further differentiation of linear and non-linear musical time into these 5 types of musical time, I will argue that it is not enough from a phenomenological point of view to distinguish between linear and non-linear as referents to causal direction if this does not account for whether or not any causality is actually perceived.

In regard to music governed by *vertical* time, Kramer writes:

> ”They lack phrases (just as they lack progression, goal direction, movement and contrasting rates of motion) because phrase endings break the temporal continuum. [...] The result is a single present stretched out into an enormous duration, a potentially infinite “now” that nonetheless feels like an instant.” (Kramer, 1988, pp. 55)

Examples of such music include Geörgy Ligeti's orchestral work, *”Atmosphêres”* (1961), Iannis Xenakis' *”Metastasis”* (1954) or his musique concrète piece, *”Concret PH”* (1958) as well as Krzysztof Penderecki's *”Threnody to the victims of Hiroshima”* (1959). Even as early as 1918, Danish composer Rued Langgaard's work, *”Sfærenes Musik”* (Music of the Spheres) shows musical traits of stasis and what could be perceived as a prolongation of the ”now”. Many more works could be mentioned by these and other composers, styles and periods. Further examination, however, of the different musical styles associated with vertical musical time will be covered in chapter 4.

Interestingly, also heavily repetitive examples of so-called minimalist music are accommodated by Kramer's vertical time category, including the work of composers like Steve Reich, La Monte Young, Terry Riley and Philip Glass. On this special form of vertical music, Kramer notes:

> “One might think of such works as purely linear, but listening to them is not a linear experience, despite their internal motion. Because in such pieces the motion is unceasing and its rate gradual and constant, and because there is no hierarchy of phrase structure, the temporality is more vertical than linear.” (Kramer, 1988, pp. 57)

Knakkergaard describes the connection between cyclic repetition and static time as follows:

> "... There [are] a number of additional conditions that allow themselves to fit under static and dynamic time as forms of understanding. These include harmonic implications when some music traditions appear static in their apparent quest for stagnation and immutability (almost timelessness) through persistent repetition of the same harmonic and rhythmic patterns in a particular mode …” (Translated from Danish, M. Knakkergaard, 2007, pp. 148)

As mentioned, in chapter 5 I propose a model for the evaluation of the degree of musical predictability. In this regard, musical pulse and repetition, which thrive in much of minimalist music, are decisive factors. I therefore find it necessary to divide this vertical, static music further. In order to illustrate the issue we must now go from the relationship between time and the concretely musical, to a somewhat higher and, for this thesis, fitting scale, the cosmic.

A cosmic perspective on repetition and irreversible change

According to the American theoretical physicist and cosmologist, Sean Carroll, events on a cosmic scale - as a function of time - can occur in fundamentally two ways: Through *repetition* or through *change* (Carroll, 2010). A simplistic definition of time could be ”what a clock measures.” He mentions three examples of how repetition makes time measurement possible: The Earth spins around its axis 365,25 times for every trip around the Sun, a pendulum in a mechanical clock sways back and forth once every second and quartz crystals vibrate 2.831.155.200 times per day. These are well-known examples to most people (perhaps with the exception of the exact number of quartz vibration per day). These three examples are patterns that, having such precision that we can predict and count on it, recur again and again in synchrony with other events that repeat just as regularly. It is repetition in the sense of a regular, cyclic expression of relative synchronization, that enables us to measure time[1]. Equally, in our own bodies there are examples of cyclic repetitive ”biological clocks” such as heartbeat and breathing.

As a consequence of the second law of thermodynamics, which states that the entropy in a closed system will always grow or remain the same (i.e. the amount of disorder always increases) (Stannard, 2013), *irreversible change* is an expression of the fact that change in the universe happens irreversibly - time moves in only one direction. Carroll refers to this principle of irreversible change as "the arrow of time", a term originally coined by British astronomer Arthur Eddington (Price, 2013). According to this principle, the future cannot be predicted, because since the degree of entropy is increasing, the number of possible futures is infinite.

I will argue that, in relation to vertically dominated music, irreversible change may be a viable metaphor for the quality of temporal forward movement that may exist in works that might otherwise be attributed a vertical musical time conception, and that occurs when the structure of such music does not let itself be predicted because the qualities of the musical structure hint towards a future with an infinite number of possible states. Irreversible change in a musical discourse, I argue, can thus be conceived of as an appendix to Kramer's *vertical time*, denoting a perceptual *non-linear* phenomenon of an *irreversibly changing vertical time* in which temporal development clearly plays a part in the musical expression but does so with no predictable goal direction.

Although it may seem extreme to turn to theoretical physics and the "arrow of time", in relation to musical perception, it is a good picture of the difference between cyclic repetition - or in other words, pulse - and non-repetitive structural[1] change. And especially Carroll's comment about cyclic repetition as a universal prerequisite for the measurement - and thus arguably *comprehensibility* - of time is interesting in this context. It is important here, I argue, to maintain a distinction between *repetition*, which may be irregular, and *cyclic repetition*, which is regular. One important difference is that the latter by virtue of its regularity allows listeners through the music to measure the passage of time. For music with a cyclic sequence of events thus applies that time is divided into limited, manageable and thus comprehensible segments unlike an irreversibly changing structure, for which time stands whole in its infinity - unlimited, unmanageable and incomprehensible.

As a consequence of the notion that musical cyclical repetition in this way can make time itself comprehensible, I find it necessary to further differentiate Kramer's concept of vertical musical time into *cyclic* and *non-cyclic vertical musical time.* And, in addition to this, append to his five musical time conceptions the notion of *irreversibly changing vertical musical time.*

A distinction can now be made between the *non-cyclic vertical time* conception associated with a piece such as La Monte Young's *"Compositions 1960 #7"* (1960) in which a single **5th** interval of **B3** and **F#4** is held for the entirety of the composition[2]; a *cyclic vertical time* that may be exemplified by Steve Reich's highly repetitive work, *"Music for 18 Musicians"* (1974-76); and an *irreversibly*

---

[1] *Structural* change because repetition can paradoxically be experienced as mutable as every recurrence may ideally be heard differently because the listener's perspective is not static but is itself changing over time due to both musical and extra-musical factors.

[2] The duration of the piece is notated as: "to be held for a long time".

changing vertical time conception, which can be associated with compositions like Iannis Xenakis' *"Metastaseis"* (1953-54), where multiple continuous glissandi in some passages of the work smoothly develop the vertical expression towards no predictable goal.

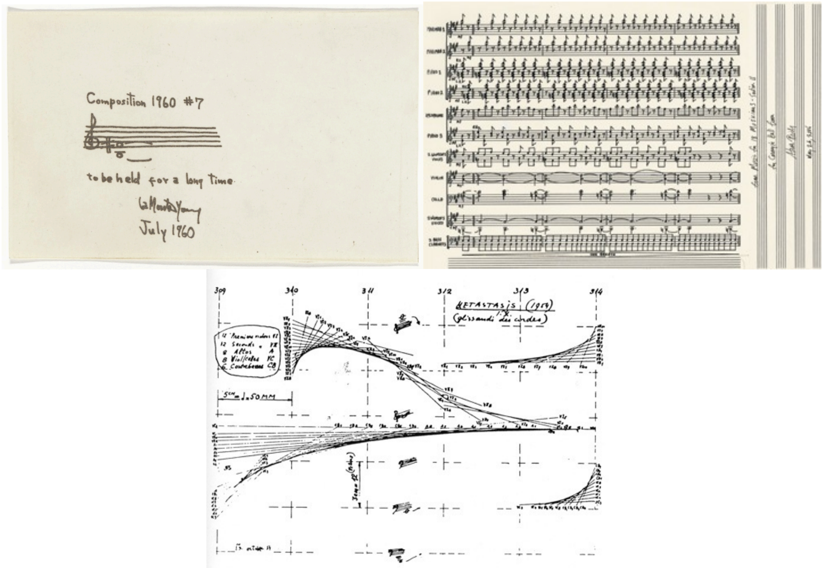


Figure 1 - Score examples of "Compositions 1960 #7", "Music for 18 Musicians" and "Metastaseis"

2.3.2.4 Auditory and musical stasis

Having established these three vertical time conceptions of music, let us turn to the question of how they relate to the notion of *stasis*. The vertical dimension can be seen as an expression of the present moment in a more absolute sense - the *now*. As such, it is an expression of a suspension of the flow of time, an ideal point with no duration of its own. It is stasis. A field of potential in which nothing can take place. It contains no process, no evolution from one state to another. The idea of stasis, however, is in itself - in its absolute sense at least - a problem. A complete standstill is an ideal, a fantasy, a theoretical state. In physics, according to the laws of thermodynamics the lowest possible temperature is referred to as "absolute zero". Even at this extreme cold state, energy is believed to still be present in the form of what is referred to as *zero-point energy* (Mehra & Rechenberg, 1999). Thus, structurally nothing in the known universe is ever completely static, completely without movement, without development over time.

But we can experience an *illusion* of stasis. At the moment of writing I am looking at a white plastic robot vacuum cleaner that sits perfectly still on the floor. I perceive it as being completely static - in a state of balance, of equilibrium. As anyone who has had experience with white consumer plastics will know however, the clean white color is slowly but surely changing into the familiar pale yellowish color associated with old computer cabinets and other plastic goods as it reacts with the air molecules around it, sunlight and so on. The point I want to make is that the experience of stasis is a matter of the scale of observation. Like watching a seemingly static forest bed only to find out it is crawling with ants as you change your spatial zoom - except in the example above we are dealing with temporality: I am not from my point of observation capable of perceiving changes in the plastic because they take place very slowly. The same could be said of the static impression one gets of the size of a mountain, the shape of a glass object etc. Only when I zoom out by remembering or in other ways preserving an impression for later scrutiny am I able to compare an earlier state with a latter and see that a change has taken place.

Likewise in the world of organized sound - be it music or sound design. I will distinguish between two different forms of *musical stasis* here that are closely tied to the newly established vertical time conceptions of cyclic and non-cyclic vertical time: One is associated with cyclic repetitive sound organizations, where the continuous repetition of, for instance, musical segments conveys a sense of status quo. Musical material can be several bars long and only when having listened to a number of repetitions the static nature of the music will reveal itself. This would be true for a number of works by minimalist Steve Reich as well as for a lot of electronic dance music, but it also applies in sound design (of e.g. a combustion engine, train sounds etc.) This type of static behavior is not associated with one immediate ”now” but with cyclic recurrences of the same ”now” or set of ”nows”. It works as the repeated exhibition of an objective segment of music or sound, which doesn’t change over time, rather than a subjective experience of time itself having stopped. This is because the regular repetitions behave like the tick-tock of a clock, conveying temporal *movement* in which the music or sound divides the otherwise incomprehensible and infinite flow of time into cognitively graspable sections. Although conveying a sense of focus on the vertical dimension by reducing horizontal activity to mere repetition, the cyclic nature is of itself a highly organized horizontal structure. I will call this form of stasis *cyclic stasis* since it is reducible to constituting cyclically recurring patterns. Its static nature is revealed only through a *broadening* of the temporal scale of observation, from the first occurrence towards the broader scope of recurrence.

The other form of stasis I will call *non-cyclic stasis* as it is not reducible into constituents of horizontal patterns. Here the experience of stasis comes across at a relatively *narrow* temporal scale of observation. Although necessarily being associated with a duration, the horizontal effects are subordinated vertical characteristics in the attention of the listener. The temporal connotation is that of time having slowed down to a complete standstill because the flow of time is not divided into regular comprehensible intervals - leaving the time dimension intact in its infinitude. Music and sound design exhibiting *non-cyclic stasis* are thus associated with a perceptual focus on one infinite auditory ”now” - the present one - in which horizontal events do not form perceivable patterns making the sequential order of these events without relevance to the experience.

In the above, I have directed my attention mainly on stasis as a meso-structural and musical phenomenon - *musical stasis* - describing the interrelation of sound objects as exhibiting a musically static expression. However, the concepts of cyclic and non-cyclic stasis reach into the micro time scale as well, what I will refer to as *auditory stasis*. Here the main implication of a distinction between cyclic and non-cyclic stasis is connected to the prolongation of the *vertical extension* - a concept, which will be defined later in the chapter denoting the duration of auditory verticality. For both musical and auditory stasis it applies that non-cyclic stasis by virtue of its patternlessness tends to inhibit gestalt-formations - an important aspect of this chapter which I will go further into later.

2.3.2.5 Functional moment, experienced moment and mental presence

One may understand the notion of stasis as a perceptual scenario in which succeeding *perceptual windows* when compared reveal no significant development or change in quality. The perceptual relationship between the whole and its constituents entails a fundamental problem; namely the question of how the succession of discreet units of perception can give rise to the experience of *continuity* - the experience of time flowing rather than jumping, so to speak, from one perceptual unit to the next. This phenomenon has been researched in a number of neuroscience studies. Based on a review of research in the field, Marc Wittmann suggests three hierarchical levels of *integration* as fundamental to this experience (Wittmann, 2011); within the range of milliseconds he calls *functional moments* those basic temporal building blocks of simultaneity in which temporally separate events are perceived as co-temporal, thus within a functional moment, no duration is perceived between events even though a duration does physically exist; a step higher up in the hierarchy, an *experienced moment* represents the integration of such functional moments into a temporal window of experienced ”nowness” of up to a few seconds in duration; lastly, Wittmann refers to *mental presence* as an experienced presence of multiple seconds in duration, which holds these *experienced moments* in short-term memory:

> “The ”experienced moment” can be related to more elementary units of perception. For example, it has been proposed that the temporal integration mechanism of around 3 s, evoking our feeling of nowness, integrates successive processing units of around 30 ms, ”functional

> moments" (Pöppel, 1997, 2009; Szelag et al., 2004)." (Wittmann, 2011, pp. 4)[3]

And further:

> "…[…] it was proposed that different mental processing stages have temporal properties matching the found temporal integration levels of 30, 300, and 3 s (Atmanspacher et al., 2004; Atmanspacher and Filk, 2010). In combining the empirical findings, temporal integration of a few seconds has been suggested to provide the logistical basis for the subjective present (Pöppel, 1978, 2009; Fraisse, 1984; Szelag et al., 2004)." (Wittmann, 2011. Pp. 4)[4]

An important aspect in the identification of these three integration levels has been an estimate of the temporal interval between events necessary for the events to be experienced as sequential rather than simultaneous:

> "The sensory systems have different temporal resolutions for the detection of successiveness or non-simultaneity. The highest temporal resolution (the lowest threshold of detection) is observed in the auditory system, where two short acoustic stimuli which are only 2–3 ms apart are detected as non-simultaneous. The visual and the tactile system have a lower temporal resolution with respect to non-simultaneity with thresholds of some tens of milliseconds; inter-modal stimulation leads to the highest thresholds (Exner, 1875; Lackner and Teuber, 1973; Kirman, 1974; Lotze et al., 1999)." (Wittmann, 2011, pp. 2)[5]

Another aspect is the border at which a temporal order of events can be perceived. This border is observed to range from 20-60 ms and appears to be universal across

[3] (Pöppel, 1997, 2009; Szelag, Kanabus, Kolodziejczyk, Kowalska, & Szuchnik, 2004)

[4] (Atmanspacher & Filk, 2010; Atmanspacher, Filk, & Römer, 2004; Fraisse, 1984; Pöppel, 1978, 2009; Szelag et al., 2004)

[5] (Exner, 1875; Kirman, 1974; Lackner & Teuber, 1973; Lotze, Wittmann, von Steinbüchel, Pöppel, & Roenneberg, 1999)

modalities - the auditory system being at the lowest end of the scale[6] (Wittmann, 2011, pp. 2).

As mentioned, one may understand the notion of stasis as a perceptual scenario in which succeeding perceptual windows when compared reveal no significant change or development. On the basis of the above this can be said to take place between the temporal windows of what is here referred to as *experienced moments*. In other words, stasis may be said to exist where nothing provokes the punctuation between *experienced moments*.

The experience of stasis may additionally be connected to a prolongation of the *experienced moment* and occur when no separable events or punctuation break points have been presented during a period of more than about 3 seconds, whereby the perceived cohesiveness between events begins to dissolve:

> "A temporal interval with duration exceeding about 3 s is experienced as being qualitatively different than shorter duration. When two events are separated by an interval of, say, 6 s, the experience of "emptiness" evolves, events are not bound together and the length of the interval separating the two becomes the focus of attention (Wackermann, 2007). A pause in a conversation, if it reaches 6 s, might be felt as disturbingly long. In that sense, duration longer than 3 s leads to the predominant experience of an extended temporal interval." (Wittmann, 2011, pp. 5)[7]

In both cases (i.e. stasis as similarity of perceptual windows, and as prolongation of *experienced moment* duration) the decisive factor is an absence of perceivable event separability.

The existence of separable events as a precondition for the applicability of the neuroscience research is obvious because clearly separable events form the basis for most of the experiments reviewed by Wittmann. The same may be said about clear onset and termination of singular events (or "filled intervals" (Wittmann, 2011, pp. 4)), which enclose a duration. Furthermore, the findings presented here can in themselves be seen as a manifestation of an epistemological approach, which is conditioned by an underlying tendency to understand the whole as a product of

---

[6] The sound *grain* associated with granular synthesis may be seen as a music-technological parallel to such perceptual windows, in effects playing on the temporal integration levels and the border of temporal order recognition proposed by the studies mentioned here. Granular synthesis in the traditional sense works with designing sound from a time scale similar to the *functional moment* suggested here. The prototype, NowEngine, on the other hand, which is presented in Appendix A, extends the granular technique to use grains of a size no longer confined to this time scale and enters the realms of the so-called experienced moment.

[7] (Wackermann, 2007)

constituting parts - treating the phenomenon of *continuation* as a side product so to speak, and not the other way around. It may further be argued that the basic notion of ”moment” and any talk of a multiplicity of separable ”moments” all together is an example of such a conditioning. The experienced continuous in this respect is seen as a reconstruction of an original continuity of the perceived environment by way of remerging discreet functional moments into an experienced whole. A ”moment”, as a separable entity unique from the continuum, may, regarded from a phenomenological or post-structuralist perspective, however, be seen as only ever existing in retrospect, or, as will be exemplified by Henri Bergson’s idea of *pure duration*, through an abstraction of time projected onto space.

2.4. DISCRETE AND CONTINUOUS SPACE-TIME

The philosophical point of departure for this thesis’ treatment of sound and music perception includes a basic distinction between the continuous and the discrete - that is, a distinction between the idea of reality as a unified continuous whole and our perception of that whole as made of separable constituents. This distinction can be made in both the temporal and the spatial domains - the latter of which will be the subject of chapter 3. Auditory verticality as a perceptual phenomenon can be said to express a cancellation of our ability to segregate auditory stimuli on the time dimension and can thus be associated with an experience of the dissolution of boundaries between individual sound structures into a continuous whole. In other words, auditory verticality is governed by an *absence of temporal interval.* This effect, as I argue in the thesis, is promoted by objective qualities of the sound structures themselves in conjunction with the subjective competence of their listener.

2.4.1. BERGSON AND THE NOTION OF *PURE DURATION*

Henri Bergson (1859-1949) wrote in *Time and Free Will* (Bergson, 1910), (Bergson & Pogson, 2001, ch. 2) about *durée pur*, pure duration, which he opposes to the quantified time of measurement. It is a subjective, intensive time associated with consciousness, as opposed to an objective, extensive time that can be analyzed. Things in space are said to be extended - that is, they extend in three dimensions. Pure duration, says Bergson, is a property of consciousness and therefore cannot be extended cartographically like for instance a calendar mapping time as we would map space. Time in this sense is not a series of moments that come one after the other but a flowing, continuous experience.

Humans, through consciousness, *endure,* and time, he says, is the experience of this qualitative *pure duration,* and as such is fundamentally different from the quantitative time conception of, for instance, the natural sciences. In Bergson’s conception of time as a duration, which is *intensive* to consciousness, time cannot be measured or counted because such an act would render it *extensive* and time would thus be projected onto *space* (Birx, 2009, ch. "Bergson, Henri (1859–1941)”).

Bergson actually uses the example of music to explain pure duration as well as his distinctions between extensive *motion* and intensive *mobility* (which I will not go further into here) as the experience of a whole phrase, instead of a succession of notes, it is an experience not composed of parts but whole:

> “Now if, finally, I retain the recollection of the preceding oscillation together with the image of the present oscillation, one of two things will happen. Either I shall set the two images side by side, and we then fall back on our first hypothesis, or I shall perceive one in the other, each permeating the other and organizing themselves like the notes of a tune, so as to form what we shall call a continuous or qualitative multiplicity with no resemblance to number.” (Bergson, 1910, pp. 105)

The notion of a ”qualitative multiplicity with no resemblance to number” (i.e. multiplicity, which is uncountable, continuous, much like e.g. the color spectrum) marks an essential point in Bergson’s philosophy and ultimately enables him to argue for the existence of free will. I shall draw a more humble use of it here, namely as an argument for the seemingly paradoxical statement that music, which is dominated by a vertical expression, may in fact lend itself to horizontal change as an expression of such an uncountable and continuous qualitative multiplicity. Bergson further states:

> “We are thus compelled to admit that we have here to do with a synthesis which is, so to speak, qualitative, a gradual organization of our successive sensations, a unity resembling that of a phrase in a melody.” (Bergson, 1910, pp. 111)

The ”phrase” here stands as a model for the experience of continuity of succession in which temporally divided stimuli (notes) may give rise, nonetheless, to a non-divisible unified experience (phrase), somewhat parallel to the conception of the present moment presented by Daniel Stern. Each new note attaches itself to the previous notes in an ongoing flow in which the present merges with past in an occurring pure duration ”with no resemblance to number.” It is, however, a different matter to use a unified experience of a phrase to exemplify a conception of time than, as I attempt here, to use Bergson’s notion of pure duration to shed light on the inner workings of music perception per se. I will argue that the music listening process may in most cases, contrary to Bergson’s use of the musical phrases to exemplify his point, be thought of as precisely an act of projecting duration into space in this Bergsonian sense. I argue that the true connection between music listening and pure duration lies less in horizontally dominated phenomena such as phrase, melody or rhythm, but may be much more easily associated with the experience of music that is vertically dominated - in which the division of time into perceivable patterns and temporally juxtaposable entities are at a minimum. These ideas will be further elaborated and incorporated in a broader model of music perception presented in chapter 5. I will, however, already here make the case that in an instance of musical listening in which past and future play a significant role in the experience of the present (e.g. when one knows the music so well that future musical events can be projected by the listener into the future based on prior experience with the musical piece) time may become *extensive* - externalizable to the listener. The listener in such a situation may for example anticipate with absolute precision that the refrain of a pop song will arrive in one and a half measures. As evident here, time is being measured and counted and thus quantified. It is laid out ”in front of” the listener by the listener and thus has *extensity* - a notion associated by Bergson not to pure duration but to *space* and time’s abstraction onto space.

The experience of continuous music or music which expresses an undivided continuum thus resembles Bergson’s description in the sense that the division of time is taken out. With no intervalic division of time, the perceptual significance of past and future decreases and the experience becomes ideally vertical - or in Bergsonian terms, music listening becomes an expression of pure duration.

2.4.2. BOULEZ, DELEUZE AND GUATTARI AND THE OPPOSITION OF SMOOTH AND STRIATED SPACE-TIME

Another illustrative music-philosophical approach to this problem can be found in the opposition of *striated* and *smooth* space-time coined by French composer Pierre Boulez in the 1950s in relation to his own music (Campbell, 2010, pp. 151) and further elaborated by philosophers Gilles Deleuze and Félix Guattari in works such as ”A Thousand Plateaus” (Deleuze, Guattari, & Massumi, 1987), where the terms are used more broadly in the philosophers’ argument for a new image of thought. Put shortly, *striated* space-time represents phenomena, which are subject to division, counting, fragmentation and juxtaposition of separable entities that can be compared and analyzed based on their temporal and spatial properties. *Smooth* space-time, on the other hand, is associated with the undivided, the uniform and continuous whole.

In ”A Thousand Plateaus” Deleuze and Guattari reference Boulez on many occasions, one of which is formulated in the following way:

> "In the simplest terms, Boulez says that in a smooth space-time one occupies without counting, whereas in a striated space-time one counts in order to occupy." (Deleuze et al., 1987, pp. 477)

And further:

> "…the striated is that which intertwines fixed and variable elements, produces an order and succession of distinct forms, and organizes

> horizontal melodic lines and vertical harmonic planes. The smooth is the continuous variation, continuous development of form; it is the fusion of harmony and melody in favor of the production of properly rhythmic values, the pure act of the drawing of a diagonal across the vertical and the horizontal." (Deleuze et al., 1987, pp. 478)

The ”fixed and variable elements” associated with striation may be connected to, for example, fixed pitches. An ”order and succession of distinct forms” could be exemplified by, for instance, musical repetition, causality and thematic development. The continuous variation and development associated with smooth space-time may be structurally represented by for example glissandi, gradual dynamic changes or the obscuring of note onsets and terminations - all of which are key components in the compositional principles of vertically dominated musical styles, as we shall go into in chapter 4. The *smooth* in the above quote is also explained on a more qualitative level as resembling Bergson’s idea of pure duration in its seemingly intensive focus on *intensity* (which is internal to consciousness) as opposed to *magnitude* (which is viewed as external to consciousness). ”Rhythmic value” and the ”pure act of drawing a diagonal across the vertical and horizontal” may be seen in this intensive light.

Boulez' original conception of smooth space-time in music may be understood to accommodate temporal interval (just as with Bergson’s example of phrase), albeit emphasizing a non-pulsed distribution of musical events. While Boulez used the terms in relation to his music, my use of them here serves the purpose of paving the way for a perhaps slightly more radical understanding of their meaning by allowing no temporal interval to occupy the smooth because the existence of interval presupposes division. My use of the terms is therefore anchored perhaps more strongly in the further elaboration of them carried out by Deleuze and Guattari. Here a number of explanatory models are used to exemplify the workings of smooth and striated space-time of which we have already discussed the musical model. I will, in the following, confine myself to highlighting a few short explanatory examples from the *technological* and *maritime* models.

From a technological perspective the difference between fabric and felt is illustrative.

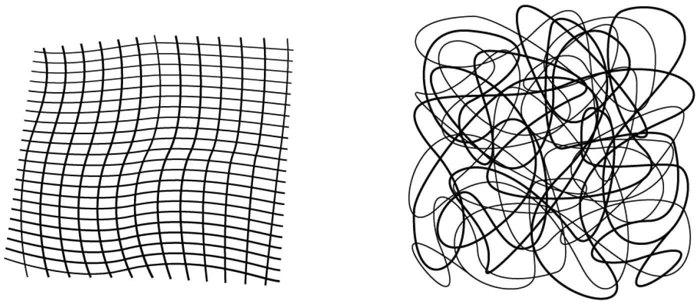


Figure 2 - Fabric and felt.

A fabric is an expression of striated space by for example having two kinds of clearly separable elements, a horizontal and a vertical thread, that intertwine and can be said to have each their own function in relation to the other. A fabric is delimited in size and by having a top and bottom. It thus represents a closed space. Felt, as an illustration of smooth space, is very different:

> "It implies no separation of threads, no intertwining, only an entanglement of fibers […] it is in principle infinite, open and unlimited in every direction; it has neither top nor bottom nor center; it does not assign fixed elements but rather distributes a continuous variation." (Deleuze et al., 1987, pp. 475)

Further, the cartographic subdivision of the sea into a grid of longitudes and latitudes provides a simple image of a striation of the smooth. And one might add the subdivision of land into national states, agricultural fields and so on. Ultimately one can speak of the striation of an ecologically uniform system of the evolutionary process of life into life forms and species, or the universal process of ever-changing matter into separable minerals, molecules, objects and so on carried out by consciousness out of necessity. A fundamental dichotomy of the holistic and the reductionist.

Smooth and striated space-time as an image of thought thus proposed by Deleuze-Guattari may be seen as an expression of ever-present perceptual and cognitive processes, which, on one hand, make the world comprehensible and graspable, ready for analysis, navigation and rational thinking, while in the process reducing the world from its unlimited continuum to a finite number of labeled and countable differing parts.

2.4.3. GESTALTS AND VERTICALITY

2.4.3.1 Gestalt principles

Similar issues concerning the relationship between part and whole may be found in the field of gestalt psychology. The term, *gestalt* (For example, Glass, Holyoak, & Santa, 1990), means shape or form. An idea of gestalt psychology is that we perceive the world as comprising meaningful and complete objects, not as a series of independent parts. These gestalt are formed on the basis of certain gestalt principles, whose visual implications serve as useful demonstrations. The picture below shows a collection of illustrative figures that demonstrate gestalt principles.

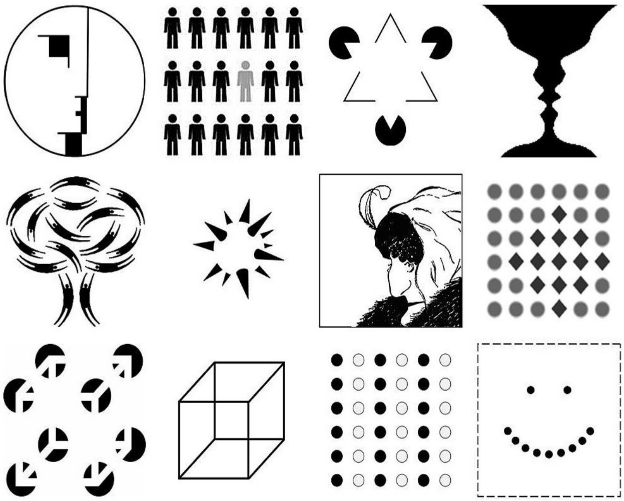


Figure 3 – Common illustrations of gestalt principles

Figure-Ground organization refers to the perceptual organization of stimuli into a clearly defined object or *figure* as opposed to a vaguely defined background. The *principle of Prägnanz* states that perceived stimuli are organized into the simplest possible experience. According to the *law of proximity,* objects that are near to one another in time or space are perceived as belonging together. Because of the *law of closure,* incomplete figures can be perceived as complete or whole. The *law of common fate* states that objects that are moving together are perceived as belonging together. According to the *law of good continuation,* objects that are aligned along a line or curve are perceived as belonging together, and we will perceive the simplest, smoothest path rather than a complex path. Symmetric objects tend to stand out as figures over a background due to the *law of symmetry*.

Another principle, which governs the perceptual grouping of all of the above, is the possibility of *perceptual set*, which can be described as a preparation for the perception of a specific gestalt through prior experience. Showing a number of complete house drawings to a person before showing a drawing where only parts of a house are shown will increase the likelihood of that person creating the gestalt of *house* when seeing the latter drawing. This principle may be associated with music by the way we are able to perceive patterns, where there strictly speaking are none, because of prior experiences with the specific pattern. This may be seen on both a horizontal and vertical level. Horizontally, for example, through the assimilation of a slightly ”untight” drum pattern into a learned schema of such a pattern in which irregularities in metrical precision are ignored. Vertically we accept, for instance, the equal-tempered tuning to be ”clean” even though it strictly speaking is false if one compares the intervals with the exact distribution of a harmonic overtone series.

2.4.3.2 Auditory Scene Analysis

Inspired by these visually derived gestalt principles, Albert Bregman in his research on Auditory Scene Analysis attempts to explain how we form auditory gestalts out of the complex continuum of sound from the many different sound sources whose sum is picked up by our ear. In relation to the auditory grouping process involved in the segregation of such complex sound stimuli, Bregman writes:

> ”If a group of components have arisen from the same physical event, they will have relationships between them that are unlikely to have occurred by chance.” (Bregman, 1994, pp. 221)

Of course, having an auditory system that lets us identify separate physical events is a practical ability from an evolutionary point of view. He further distinguishes between what he calls *sequential* and *simultaneous organization* when describing the human brain’s ability to separate incoming auditory stimuli into individual sound sources (Bregman, 2005, pp. 37-38). Both sequential and simultaneous organization deal with the problem of segregating the occurring auditory ”now” - the sounding simultaneity or verticality - into its constituting sound sources. The difference is whether the method applied can be said to focus on horizontally based cues (sequential) or vertically based cues (simultaneous).

He further uses the notion of *auditory objects,* to refer to single sounds that are perceived as one unit and *auditory streams* to describe the percept of sequences of sound that unfold over time. The research shows that the sound we hear is grouped into *auditory objects* and *streams* depending on the degree of *similarity* in both the time and frequency dimensions.

One could elaborate, based on Bregman’s categorization, that, although both auditory objects and streams successfully form gestalts via simultaneous and sequential organization methods respectively, there is a fundamental difference to how it happens. While auditory objects are truncated from the complex continuum of sound on both the vertical and horizontal dimensions, auditory streams, if they are to maintain their function as a stream, are only segregated vertically. In other words, auditory objects are delimited by time and space, while auditory streams have primarily the space dimension as the basis for gestalt formation.

For sequential organization the auditory assessment of similarities is based on *pitch, timbre, spectral properties, temporal properties, location in space* and *intensity*. Interestingly, according to Bregman's research the musical phenomena of rhythm and melody tend to arise only within the same auditory stream. The same applies for *"...fine details of temporal order..."* (Bregman, 1994, pp. 143)

For simultaneous organization, the grouping relies on *synchronized changes (or common fate), common origin in space, "old-plus-new" heuristics* and *spectral relations* such as identifying *multiples of a common lower frequency.*

Old plus new heuristic refers to situations where acoustic components can plausibly be interpreted as continuations of a sound that has just occurred and is based on the following rule: *"If you can plausibly interpret any part of a current group of acoustic components as continuation of a sound that just occurred, do so and remove it from the mixture."* (Bregman, 1994, pp. 222) This method can be criticized for actually not being an analysis of simultaneity but temporality since recollection of prior stimuli is used for the segregation of current sound components.

Spectral relations entails the fact that harmonically related partials are grouped based on *harmonicity* - in other words, partials that share a common lower frequency are perceived as stemming from that lower fundamental frequency and thus plausibly from the same sound source. (Bregman, 1994, pp. 227)

The principal of *common fate* is used by our auditory system to organize into a single auditory stream or object sounds for which *"...different parts of the spectrum change in the same way at the same time."* (Bregman, 1994, pp. 249). The changes referred to by Bregman exist in two types: frequency modulations (FM) and amplitude modulations (AM). Correlated changes in amplitude and frequency, whether they are obvious or subtle, have a tendency to enhance the perceptual grouping of those sound components as the auditory system assesses they must be from the same acoustic source.

In some of Ligeti's music the composer goes to great lengths to avoid the perception of common fate in regard to musical meso-structure. His agenda in these efforts is the avoidance of segregation that might come about where several voices, in having a common fate, may be grouped together and as such stand out from the surrounding musical texture. The result being an inability by the listener to segregate meaningful musical individualities from the continuous flow of sound.

Spatial correspondence refers to how spatial proximity can lead to grouping. Tones that are close to each other in pitch tend to group together to form the same auditory stream, whereas tones that are far apart in pitch tend to be perceived as belonging to each their own auditory stream.

Auditory Scene Analysis hierarchies

In this way hierarchies of groupings form. In, for instance, a concert setting the sound of the symphony orchestra as a whole is grouped separately from that of the audience sounds. Within the orchestra stream the oboe stands apart from the violins and so on. Auditory stream segregation is an expression of a perceptual *striation* of the *smooth* auditory continuum encountered by our eardrum. It seems to be implied by Bregman's research that the listener has little or no control over this process. A composer or sound designer, on the other hand, who is responsible for the sound heard can organize these sounds in a manner that either eases or challenges our ability for auditory stream segregation. Vertical dominance can be seen as a result of horizontal submission. It is what occurs when, for some reason or other, our ability to perform the horizontal striation is challenged to such an extent that our attention directs its primary focus on the sonorous simultaneity.

Auditory Scene Analysis and musical meaning

Horizontal striation is in some ways a prerequisite for *meaning*. This is true for language where intelligible separation of words and sentences is necessary in order to understand the message they carry, and it is true for musical meaning as well, which traditionally also emerges out of striation - that is, tones striate pitch-space by means of pitch and timbre, and striates time by having an attack, a duration and a termination. Polyphony, rhythm and harmony form hierarchies and causal movement - a fabric of meso-structure often referred to as musical texture. As striation is lifted we come closer to the continuum - to smooth space-time, as Boulez puts it - and the notion of musical meaning may be argued to shift to a different semiotic resource. More on this in chapter 6.

Auditory Scene Analysis in rhythm, melody and dissonance

According to Albert Bregman, *rhythm, melody* and *dissonance* are all phenomena that are dependent on their constituting components being in the same auditory stream in order to emerge in our perception. If, say, the notes of a melody are separated according to the principles of auditory scene analysis so that they are no longer part of the same auditory stream, the phenomenon of *melody* would disappear, and only fragmented individual notes would be perceived. The same applies to the auditory events involved with the perception of rhythm and dissonant intervals.

Striation as the creation of gestalts

The dichotomy of smooth and striated space-time is precisely an expression of gestalt creation. In sound and music, because it is based in time as well as space, the spatial continuum, out of which visual objects are formed in the often portrayed examples of gestalt principles, does not stand alone but is accompanied by a temporal continuum. The horizontal auditory stream segregation or *sequential* organization striates time, while *simultaneous* organization striates spectral space.

Do we ever stop striating?

Since human perception is so tied to, and governed by, this gestalt grouping mechanism in order to comprehend, one could ask whether this perceptual principle ever ceases to function? Do we ever not group or segregate? Do we ever step into conscious states of non-division, of a non-duality in regard to the world around us - and as an absolute consequence in regard to also ourselves as connected to that world? The gestalt division of the space-time continuum can be seen as ”a way our minds work”, as an illusion and a very practical one. I do not mean to say that the universe is fundamentally chaotic. Structural regularities are everywhere around us, but they represent organization in a continuous flux, a process. This brings us back to the illusion of stasis. Experienced stasis in this respect as a perceptual illusion bound to gestalt division attempts to fixate an object - be it auditory, visual or material - in space and time. The interplay of different musical gestalts into larger meso-textures of a musical fabric narrates a meaningful process. The robot vacuum cleaner forms a gestalt as we perceive its physical properties as different from the air around it and the floor beneath it. In any practical sense it is in fact, to a human, different from its surroundings. But in the larger scope it is so only temporarily. On a very small spatial scale or a very large temporal scale it becomes obvious that it is interconnected to the ever-changing flux of material stages - a temporary lump in the universal gravy of space-time. It is fully integrated with the flow of everything. It is an accumulation of energy that changes very slowly. Objects seem static to us when we do not perceive any change within the time of observation.

2.4.3.3 Perceptual windows and gestalts

The idea of *perceptual windows* may be associated with whether or not stimuli received by the auditory system can be perceived as having changed over time. The idea being that a sufficient difference can be detected between the qualities of sequential perceptual windows and that this may lead to an experience of development rather than stasis. I will argue that the experience of development per se cannot be said to promote horizontally oriented gestalt segregation. In regard to VDM this implies that the musical flow does not lend itself to intuitively being subdivided into musical sound object time scale gestalts. The musical substructure, which in an HDM context may be clearly segregated, tends to act in VDM as an intact whole and does not lose this continuity even during temporal development. While the perception of development must be said to implicate a comparison between the present state of the music and a former one in order to discover a change having happened (thus suggesting a juxtaposition of perceptual windows, but not necessarily the existence of separate gestalts), VDM is characterized by a reluctance towards presenting cues for such division of substructure beyond the micro time scale (as well as the *vertical extension* time scale covered shortly). Long continuous glissandi for example (a key compositional concept in e.g. Xenakis' *"Metastaseis"*), while being clearly perceived as developing, do so in a smooth, unstriated movement. Continuous developments such as glissandi are clearly not static and thus, arguably are not an expression of auditory verticality although they do not promote the formation of separable gestalts but on the contrary are experienced as one unified movement. This is a clear illustration of how essential the differentiation is between the notion of a perceptually static vertical *expression*, which in the case of a glissandi is continuously changing over time primarily on the pitch parameter, and the notion of vertical *domination* in which the expressive attributes associated with auditory verticality (e.g. harmony, timbre and micro-texture) receive a comparably higher degree of significance due to an absence or obscuring of horizontally biased structures (such as pulse, rhythm and causal directionality). *"Metastaseis"*, and glissandi per se, can be regarded as vertically *dominated*, while they do not, on the other hand, represent a purely vertical *expression*.

The rather complex issue of the relationship between causal direction, anticipation, the formation of gestalts and vertical or horizontal domination is further elaborated in chapter 5.

2.5. CONCLUSION

Musical verticality, as a compositional phenomenon in western music tradition, has extended from residing in pitch-space as a result of polyphony to occupy the *wide open sound world* of spectral space, thus allowing structures with complex spectral properties to form "sound objects" and "spectromorphologies".

There is no sound without duration. The horizontal dimension of time is ever-present in any manifestation. Therefore, in an absolute sense, there can be no separable, sounding verticality. We can make snapshots of the measured value of an instant in time - a *sample* - but this sample is not sound. Were we to repeat the same sample at say 44.100 times per second, there would be silence. From a physical perspective, sound is pressure variations in a medium over time and thus, an inherent horizontality is present out of necessity. But at a certain point a shift occurs in perception. It is a shift between hearing sequential occurrences of separate events to hearing unity - as is the case with the perception of pitch.

Certain quantifiable durations are significant in the formation of auditory verticality. 50 milliseconds is the duration of one period of a 20Hz sound wave and roughly represents a point of transition between two very different auditory phenomena - rhythm and pitch. The cognitive *present moment* can be regarded to exist in a temporal frame of 2-5 seconds. *Basic units of perception* within the area of 20-150 ms are grouped into larger meaningful units by our brain - suggesting a form of verticality.

I will suggest, however, that *auditory verticality* cannot immediately be defined as a fixed duration of a particular number of seconds or milliseconds. Nor should it be allowed to contain any obvious *horizontal* musical traits or groupings of meaningful structures associated with the meso time scale, such as motives and phrases. And it should not, for that matter, be viewed as a cognitive process reacting on musical expectations caused by horizontal cause-and-effect relations (that would be a *horizontally* focused musical ”now”.) Its duration is relative to, for instance tempo and activity in the preceding music as well as to any audiovisual or interactive context the music co-constitutes. To define *auditory verticality* - if argued to exist at all - must therefore necessarily be done on the basis of a phenomenological consideration.

Let us first say, phenomenologically speaking, that *auditory verticality* exists on its own right, autonomous from horizontality. Let us then say that it does so within a temporal window. As we have observed, the duration of the temporal window of auditory verticality does not confine itself to Wittmann’s lowest threshold of detection at 2-3 ms for auditory stimuli, nor does it let itself be defined as Stern’s ”present moment. ” The concept of ”sound”, as it is used by Brolinson and Larsen to describe properties of popular music, includes some spectral and dynamic characteristics of production, but allows also musical meso time scale structures as co-constituents - inviting horizontally dominated structures inside the fold and therefore it is not suitable as an expression of verticality.

The notion of ”vertical musical time” put forth by Kramer applies to repetitive music as well as irreducibly changing music and thus striates the dimension of time. This led me to differentiate Kramer’s concept of *vertical time* into *cyclic* and *non-cyclic vertical time* governed by *cyclic* and *non-cyclic musical stasis* respectively. Additionally, I found it necessary to append to these musical time conceptions *irreversibly changing vertical time* as a musical time conception in its own right in which temporal development clearly plays a part in the musical expression, but does so with no predictable goal direction.

I propose that in regard to auditory verticality, as far as the horizontal dimension is concerned none of the covered principles of gestalt theory and auditory scene analysis must be allowed to act within the percept - including, for instance, the *law of continuity*, which is relevant in regard to, for example, glissandi movements, the *law of common fate* and so on. Auditory verticality exists as a percept only as a coherent, *unstriated*, *unquantified* whole with no beginning and no end. In the experience of such a whole, the listener does not *segregate sequentially* the continuous flow of sound. (The structural implications of auditory verticality are covered in the next chapter.)

Summing up the covered ground in this chapter, I thus suggest two primary concepts that will allow for a definition of auditory verticality suitable for the further investigations into vertically dominated music in the thesis: the *vertical extension* and the *vertical expression*.

2.5.1. VERTICAL EXTENSION

As discussed, absolute verticality is a mathematical ideal of a line in space with no duration - an expression of a practically impossible separation of the spatial and temporal dimensions. As further discussed, from a perceptual point of view auditory verticality does exist. I propose to address the temporal delimitation of the perceptual phenomenon of auditory verticality as a temporal extension of this vertical line - a *vertical extension*[8]. The duration of this extension is flexible and its defining factors are qualitative. The vertical extension defines a temporal window within which perception can be described as an expression of Bergson's *pure duration* and Boulez' *smooth space-time.* It claims this status by refusing temporal subdivision - the main threat to verticality - into separate gestalts, or sequential auditory groupings. Its constituting sound material is continuous rather than discrete and promotes an illusion of *auditory stasis.* Additionally, the vertical extension may be prolonged when this material is governed by *non-cyclic stasis* due to this stasis type's patternlessness and thus inhibiting effect on gestalt-formation.

(To get a hands on sense of the perceptual difference between cyclic and non-cyclic auditory stasis please refer to Appendix A covering the prototype, NowEngine. Cyclic and non-cyclic stasis is described in the section about the [Range] parameter.)

This temporal window of auditory verticality could be relatively short or it could be upheld for a considerable amount of time by means of compositional principles that have this deliberate aim - this will become more evident from the following chapters.

In this way the vertical extension may exceed the micro time scale and scatter the stringency of Curtis Roads' time scale model. In such cases it could be argued to function as a prolongation of the *micro* time scale into extended durations, refraining from entering sound object-territory by omitting the music-structural characteristics that facilitate individuality of sound object events and therefore also any formation of such events into higher-level phrase structures. It is thus useful within the discourse of this thesis to extend Roads' time scale with an additional time scale - the *vertical extension time scale* - which will play a central role throughout the thesis. The vertical extension time scale encompasses the micro time scale but may likely exceed it. Since the vertical extension only reaches as far as to the beginning of any perceivable gestalt-forming change of the vertical expression occupying it, and since the sound object time scale allows for internal movement (a simple example of which would be a onset-continuation-termination envelope) the vertical extension time scale situates itself between the micro- and the sound object time scales in Roads' model as follows:

[8] The direction of the extension is really horizontal, temporal. The term defined here refers to a temporal extension of the vertical line into having duration.

- Infinite
- Supra
- Macro
- Meso
- Sound object
- ***Vertical extension***
- Micro
- Sample
- Sub sample
- Infinitesimal

2.5.2. VERTICAL EXPRESSION

While the notion of *verticality* implies ”no time”, within this phenomenologically defined vertical extension, micro-textural activity *is* possible. An experience of verticality can be maintained even with internal micro time scale movement. I will refer to the vertical extension’s internal expressive structure as its *vertical expression* - a concept constituted by harmonic, timbral and textural properties, which shall be examined further in the next chapter on the structure of auditory verticality.

2.5.3. AUDITORY VERTICALITY

Based on these considerations I will define *auditory verticality* as:

> **A perceptual phenomenon, which is facilitated by an illusion of ”auditory stasis” caused by structural properties of a ”vertical expression” and occurring within a temporal window of a ”vertical extension”.**

Vertically Dominated Music (VDM) can subsequently be expressed as:

> **Music in which the expressivity of auditory verticality dominates over that of auditory horizontality.**

A vertically dominated expression can therefore occur when sound object and meso-structure divisions are diffused - when onset and termination of an onset-continuation-termination envelope are so subtle that the separation between the structures’ individual constituting parts becomes perceptually obscured. The listener’s ability to segregate is diminished and the experience is of a single sound-mass in which individual events tend to disappear into the larger whole of which they are a part. Further elaboration of the music-structural implications of VDM is presented in chapter 4.

2.5.3.1 Auditory horizontality

Based on this definition of *auditory verticality,* let us refer to *auditory horizontality* as:

> **A perceptual phenomenon of those auditory developments that exceed the ”vertical extension” and which facilitate the formation of clear temporally truncated gestalts.**

These gestalts in turn make possible the perceptual emergence of horizontal musical phenomena such as pulse, rhythm, tempo and groove as well as horizontally dominated structures such as melody, themes and thematic development, harmonic progression, causal direction, imitation and repetition.

When speaking of horizontality in regard to sound and music I will again distinguish between structural and phenomenological implications. I make the distinction here in order to reiterate in the name of clarity what has already been discussed several times, namely the rhetorically problematic paradox that on the structural level and in the context of the physical reality we live in - including in the context of sound - what we perceive as vertical in fact *has* duration. And that this does not exclude a sense of simultaneity in the listener.

The *phenomenological* dimension of auditory horizontality can be said to start, so to speak, where the temporal window of auditory verticality ends - in other words, where the illusion of auditory stasis is broken. It is associated with what Boulez (as well as Deleuze and Guattari) refer to as *striated space-time.* Its stimuli are perceptually organized through what Albert Bregman calls *sequential* rather than *simultaneous auditory segregation*, it exists as an expression of Bergson’s conception of *quantified time,* and it is occupied by structural manifestations at time scales greater than the micro time scale as these are defined by Curtis Roads as well as beyond the time scale of the recently defined *vertical extension*.

From a *structural* perspective, horizontal development does, however, exist at the micro time scale and within the vertical extension. Structurally speaking, horizontal movement is present within our phenomenologically defined vertical extension as it co-constitutes properties of the *vertical expression* such as pitch, timbre and micro-texture.

Based on the above and as the dichotomous opposition to VDM we can define Horizontally Dominated Music (HDM) as:

> **Music in which the expressivity of auditory horizontality dominates over that of auditory verticality.**

2.5.4. VDM AND HDM

A simplistic illustration of the relationship between VDM and HDM could look as follows:

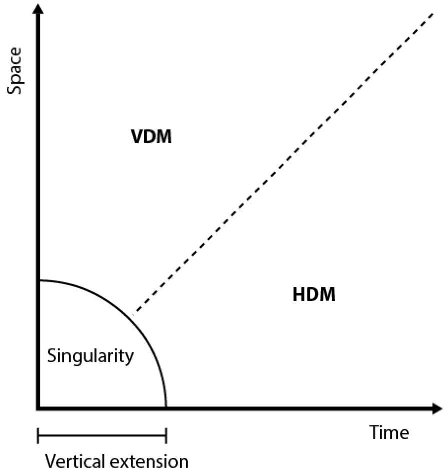


Figure 4 - VDM and HDM, vertical extension and auditory singularity in the space-time continuum

The figure above depicts the time-space continuum occupied by musical structure, the perception of which ranges from vertically to horizontally dominated. What happens at the perceptual level within the duration of the vertical extension can be regarded as an experienced fusion of temporal and spatial properties into an auditory *singularity*[9] in which horizontal and vertical structures are not perceptually separated. The joint effort of horizontal and vertical structure forms here an auditory expression that is at the same time pure form and pure duration. The concept of auditory singularity is taken up as part of a broader model of the relationship between musical structure and perception in chapter 5, which also places vertical and horizontal domination in relation to musical anticipation and proposes a perspective on the definition of "music" based on these ideas.

The internal organization of this singularity - that is the internal structure of the vertical expression - is the subject matter of the next chapter, which looks into a number of striations paradigms that have been associated with the vertical dimension in music. These include tone systems and the concepts of harmony, timbre and texture.

[9] The term is borrowed from astrophysics and the study of black holes.

CHAPTER 3. STRUCTURE OF AUDITORY VERTICALITY

3.1. INTRODUCTION

Until now we have looked at the duration of the vertical as well as its philosophical and physical implications - of which not least the relationship between continuity and discontinuity plays a significant role. While verticality in its absolute sense is a theoretical ideal, a phenomenological perspective allows the possibility of defining a temporal window for it to exist in. Having said that this duration is highly variable and dependent on subjective experiences of *auditory stasis* and an absence of horizontally truncated gestalts, it is now time to examine the phenomenon of musical and auditory verticality closer - this time looking at its internal structural organization.

How do we begin to describe, categorize and organize structures within this definition of the auditory vertical? When speaking of relations between fixed pitches alone the scope of the challenge is finite - however, it is not necessarily simple. When considering verticality as a spectral space of theoretically unlimited timbral resolution, the task is much more complex. If we further add to that the infinite number of possible micro-textures that may occupy a vertical extension the possible vertical expressions are - quite obviously - endless.

The main purpose of this chapter is therefore to identify a range of *vertical parameters* for VDM composition that will allow us to navigate in this smooth-spaced ocean of possibilities. The chapter also includes a number of existing *vertical striation paradigms* potentially forming a pallet of vertical subdivision concepts, which can be used as a basic reservoir for VDM composition.

The expressive structure, which occupies the temporal window of what we in the previous chapter called the *vertical extension,* can be understood as a three-part hybrid concept of harmony, timbre and micro-texture that forms a more or less complex emergent entity: a *vertical expression.* The vertical expression can be subdivided into a range of vertical parameters that may be manipulated individually. These variables cannot, however, be understood entirely as separate, and developments in one category will likely affect the others. The vertical parameters identified in the present chapter are meant to serve an adaptive compositional process. The particular striation paradigm they together represent, and which I put forth here - although undoubtedly based on crucial aspects of auditory verticality - should not be understood as an attempt to create a comprehensive framework for understanding vertical structure. The vertical expression must be regarded as a phenomenon with infinite structural possibilities -

possibilities that only increase as new technologies open new methods for sound generation and manipulation.

Qualitative terminologies based on metaphors such as bright, harsh, dull, dark etc. are not precise enough in the context of generative VDM composition, which requires concrete parameters for its formalization. A purely quantitative perspective on the other hand, which categorizes for example the auditory spectrum into specific frequency areas or specifies a frequency range in Hertz, seems to not being able to account for the way we experience sound. The perspective I take in this chapter is thus phenomenological but at the same governed by an interest in the structural and physical aspect of sound.

As will be evident, the formation of vertical structure has historically relied on a continuous refinement of various systems to build a bridge between the smooth and continuous physical world. This refinement is fueled not least by a need for a striation paradigm of the vertical auditory dimension that is capable of supporting and facilitating our ever-evolving musical expressivity.

In this chapter we will discuss a number of aspects fundamental to the structure of the vertical expression. These include such vertical striation paradigms as tone systems, temperaments and the concepts of harmony, timbre and texture. As the latter three aspects can be subject to a degree of ambiguity, their inclusion in this chapter will serve to clarify what exactly is meant by them in this thesis, as well as provide vertical parameters that can be used for adaptive sound manipulation. Furthermore, auditory verticality may be seen to also encompass acoustic properties of sound. I go briefly into this aspect of auditory verticality, but maintain its function as more of a medium for vertical expressions to sound in, rather than as part of the definition of the vertical expression per se.

These elaborations are meant to pave the way for a more precise understanding of the notion of *vertical expression.* This chapter can thus to a large extent be regarded as a definition of this term based on the components discussed. As such, the vertical expression can be understood as *an emergent hybrid concept of pitch, harmony, timbre and micro-texture.* Moreover, the chapter's role in the thesis is not least to provide a list of vertical parameters for the generation of adaptive VDM in games. This list is grouped into the discussed components, tone system, harmony, timbre and micro-texture.

The first of these components I will discuss is the striation of smooth spectral space into pitch-space by means of different *tone systems*.

3.2. TONE SYSTEMS

Throughout European music history, the problem of finding suitable striation paradigms for the vertical musical dimension has undergone several stages of development and has found a multitude of solutions. At the very foundation of such endeavors lie the different tone systems that have been implemented as the basic grid for composition with fixed pitches. The development of such schemes for subdividing the otherwise smooth vertical continuum has, in western music culture, traditionally utilized two main principles: an *identity interval,* which denotes the interval at which two different pitches are to be understood as identical (in the European music tradition an octave is the identity interval) and one or more *generator intervals* (Egeland, Finn in M. G. Knakkergaard, Finn, 2003, pp. 1641) by which the further striation of the vertical continuum can be performed. The different approaches carry their own advantages and disadvantages in regard to their degree of consonance and musical flexibility. Although many more exist I will confine myself here to mentioning a few of the more prominent examples.

The *Pythagorean* tone system was prevalent in western culture from the Greek antiquity until the 15th century. In the 12 tone Pythagorean tuning octaves (fundamental string length divided by 2) are considered identical. All 12 subdivisions are generated based on the **pure 5th** interval calculated as 3:2 of the fundamental string length. Intermediate intervals within the octave are thus accomplished by stacking **pure 5ths** on top of each other, so to speak, and lowering these by the appropriate number of octaves until a chromatic scale of 12 semitones appears with slightly varying mutual intervalic relations. The Pythagorean tone system is not circular in the sense that 12 **pure 5ths** stacked on top of each other add up to a slightly higher frequency than 7 octaves (the closest octave interval). This deviation of 1/4 semitone (or 23 cents) is referred to as the *Pythagorean comma*. Additionally, this striation paradigm renders the **major 3rd** sufficiently out of tune relative to the fundamental to be considered dissonant in contemporary music theory.

By the late 15th century the *pure* tuning scheme, while not being exactly pure, attempts to rectify the latter problem by introducing - in addition to the **pure 5th** - the **pure major 3rd** as generator interval by the proportion of 5:4 of the fundamental string length. The **major 3rd** thereby gradually became consonant in contemporary music theory and practice - a significant precondition for the further development of western musical tradition. The downside to this approach was that certain of the **5th** intervals in the resulting 12 tone chromaticism became significantly false to the ear due to their odd intervallic proportions.

Hence came the idea of the *tempered* tone systems, which abandoned the principle of generating intervals based on pure consonance and instead took the pragmatic approach of seeking to increase flexibility of modulation between musical keys. The first of the tempered tunings was the *meantone temperament*. Here all **major 3rds** are pure while all **5ths** are diminished by what is referred to as a quarter of a *syntonic comma* - distributing the offset on four **5th** intervals instead of - as in the *pure* tone system - on only one, making the interval deviation more tolerable.

The *equal-tempered* tone system, which can no doubt be considered a standard today in western music culture, is truly circular. All 12 semitones are here equally spaced leaving the **5ths** slightly flat and the **major 3rds** a bit high but allowing absolute flexibility and uniformity across keys within this 12 tone chromatic vertical striation paradigm. Although records exist of its invention as far back as 15th and 16th century (not least as a design strategy for fretted instruments), even in the 19th century the equal tempered tuning was often avoided due to its particular intonation compromises in favor of older and in some contexts more consonant tone system predecessors (Gravesen, 2003, pp. 1846). Its great potential for musical flexibility and expressivity, however, which has allowed it to prevail until today, was already in its cradle exploited by J.S Bach in Das Wohltemperierte Klavier.

These tone systems have been invented. It would be wrong to say, however, that in the striation paradigms of the vertical nothing is a given by nature. The correlation between overtones that led to the perception of consonance (the concept of *consonance* will be more thoroughly discussed later in this section) lies at the foundation of all of the above mentioned tone systems. But the striation attempts have proven to represent not only qualities of nature, but to a large extent, also culturally determined compromises between flexibility and perceived pure consonance - fueled by a wish to modulate freely between keys. The subdivision of 12 semitones to the octave is another example of a culturally conditioned standard, which is not nature given. A subdivision of 19 semitones to the octave, for example, would lead to slightly cleaner intervals. Several other approaches exist in ethnomusicology in which different subdivisions have taken root such as the South Indian 22-tone system (Potty & Parameswaran, 2015). In western tradition, micro-tonal systems divide the octave into for instance 24 quarter tones or 48 eighth tones, usually conducted in an equal-tempered manner. Also worth mentioning is the American 20th century composer, Harry Partch, who devised his own tone system of 43 tone per octave (Mitchell, 1983). With the spectral music of the early 1970s a direction was set to compose within even the open spectral continuum inspired by the natural overtone series - this was carried out, however, through the compromise of *orchestral synthesis*, which is covered later in the present chapter.

3.3. HARMONY

The term, *harmony,* has a number of connotations. In antiquity simple proportions such as 1:2, 2:3 and 3:4 were considered to be harmonic in music as well as in other arts such as architecture. Harmony here refers to an accordance with a cosmic order of the Universe that was sought after in the arts. In music theory of the European tradition after the 16th century, a harmony may refer to a chord of three or more notes, and bear the additional connotation of this chord's functional relations, tensions and directionality. On a more mundane level, harmony may also take the meaning of something that simply sounds aesthetically pleasing (M. G. Knakkergaard, Finn, 2003, ”Harmoni”).

The term, as it is used in this thesis, is associated not with particular ideals for horizontal harmonic tensions and progressions. Nor is it to be understood as in any way coupled to an accomplishment of consonance or other aims to hierarchically categorizing vertical pitch combinations based on aesthetic preconditions of pleasant sounding intervals. On a structural level, I shall regard harmony primarily as the relations between the first partials of simultaneous sounding tones.

Traditionally, harmonic analysis deals with such relations between simultaneous sounding pitches but also with the tensions, hierarchies and horizontal connections between these vertical constructs. In the tonal tradition after the implementation of the equal-tempered tone system the grid within which these chords were conceived consisted of 12 equally spaced tones per octave organized into diatonic scale modes based on a keynote. The records of what is now commonly referred to as *church modes* (lydian, phrygian, dorian etc.) originate, as far as we know, from Greek antiquity (M. G. Knakkergaard, Finn, 2003, ”Harmoni”). These diatonic modes having prevailed to the present have been accompanied over time by a large number of other scales within the 12-tone chromaticism. These include the Neapolitan minor, symmetrical, Prometheus, whole tone, and overtone scales, to name a few, and were no doubt predated by various pentatonic scales.

3.3.1. TONAL KEY

This vast repertoire of scale material has historically represented an important resource for the composition of vertical expressions and can do so in VDM as well. A detachment from tonal hierarchy and especially the absence of directness and causality, which plays a significant role in the formation of a vertically dominated musical expression, may however to some degree deprive some of these scales of their raison d'être. There is really no difference between the seven church modes if one takes away the tonal center and hierarchy between notes. Nevertheless, even in scenarios that do not as such contain a tonal key, VDM may be ”anchored” to a particular static drone - either in the form of a single tone, a harmony or a cluster - by which a point of departure is offered to the presented scales. In such cases it can make sense to speak about the music being anchored in a certain key or keynote although this anchor may be diffuse in comparison with a musical key in the traditional sense. (This anchoring can be associated with what Denis Smalley refers to as *relative pitch* covered later in this chapter).

A musical key signature traditionally points to a fundamental pitch and an associated scale. It likewise entails tonality and to some extent therefore progression and hierarchy, thus bearing connotations associated with horizontal dominance. The use of the term should, however, not be ruled out completely in the context of VDM in the sense that it presents a well established and, for musicians and composers, recognizable framework for the categorization of triad-based harmonies. Furthermore, in regard to adaptive VDM's narrative and functional coupling to elements in a computer game such as characters or locations, this framework of the musical key also offers opportunities for bi-tonal and poly-tonal composition (used by, for instance, Igor Stravinsky and Béla Bartok) in which more than one key center function simultaneously - often characterized by being as far apart as possible in the circle of fifths in order to achieve maximum separation between tonalities. The use of two or more simultaneous sounding musical keys in adaptive VDM may have significant narrative potential if each key is associated with a different in-game element.

3.3.2. HARMONIC BASE INTERVAL

Harmony may be based on **3rds** as in the tonal system, but harmonies are also built on, for example, **2nds**, **4ths** or **5ths** - which could be called different harmonic base intervals - each with its own unique vertical flavor and expressivity. (The algorithmic composition program, CALMUS, which is presented as part of a case study in Appendix B, offers the choice of such base intervals as a parameter to manipulate the ongoing compositional process in real time.) As with the previously mentioned aspects of harmony described above, the choice of harmonic base interval may be utilized narratively in its coupling to elements in the game world, lending itself as an obvious vertical parameter in a system for generating adaptive VDM in games.

3.3.3. TONE CLUSTERS

Also the use of tone clusters consisting of **minor** or **major 2nds** in close proximity is an important harmonic building scheme in VDM as will be exemplified later in chapter 4 as associated with composers such as Ligeti, Penderecki and Xenakis. In regard to the harmonic architecture of clusters the most significant trait perhaps is their register in terms of position and range. Although a tone cluster may be seen as a group of three or more simultaneously sounding notes spaced with no more than a **major 2nd** interval (e.g. Persichetti, 1961, ch. 6), it can be argued that larger clusters may contain internal "holes" in them of larger intervals, while still being perceived as one cohesive cluster - or in other terms, as one auditory stream. The number of auditory streams involved in the perception of a particular vertical expression may itself be used as a parameter for VDM composition and exploited accordingly as an expressive and narrative resource. The tone cluster is understood here as a purely vertical phenomenon. It can also, as is the case in the Polish musicologist Chominsky's theory of *sonoristics*, be conceived as being much like a melody, encompassing both the cluster's vertical and horizontal design (Granat, 2009). In the framework of Denis Smalley's spectromorphology, clusters represent *relative pitch*. Relative pitch relations are heard with less precision and instead form somewhat diffuse note-collectives. This happens if notes are stacked closely together in clusters or changed to a different pitch very quickly so the duration of the individual pitches is so short that we cannot perceive them precisely before a new pitch takes over. In what Smalley calls *intervalic pitch* relations the notes are clearly separable by the listener. In the discourse of auditory stream analysis covered in the previous chapter, relative pitch corresponds to the auditory grouping described by Bregman to occur due to the gestalt principle of proximity, while intervallic pitch perception has a stronger tendency to provoke the perception of several individual auditory streams.

3.3.4. CONSONANCE AND DISSONANCE

Another phenomenon with historically great significance within the subject of harmony is the dichotomy of *consonance* and *dissonance* and the border between the two has been subject to much change and debate. While the older tone systems such as the Pythagorean and pure tunings would exhibit obvious differences in regard to how clean or false certain notes would be perceived in the different modes, with the introduction of equal-tempered tuning, these differences have been evened out - making the distinction less overt. A degree of ambiguity exists also for the notions of consonance and dissonance. Intervals between pitches may be considered consonant or dissonant based on primarily three different perspectives: a sensory, musical and psychological perspective. The *sensory* approach explains consonance as an absence of *beats*[10] (and the degree of dissonance by the degree to which the interval produces audible beats). This absence of beats is caused by either coinciding harmonics or by having sufficient spacing between pitches to render the beats weak in amplitude (Helmholtz, 1877; Plomp & Levelt, 1965). Sensory consonance and dissonance also have a different explanation, namely in terms of *harmonicity.* According to this when the proportionate relationship between pitches resembles those of the natural harmonic series (1:2, 2:3, 3.4 etc.) their harmonic relationship will be experienced as consonant. When it does not, varying degrees of dissonance arise. Recent research suggests that between these two harmonicity may have a stronger impact on the experience of consonance than the absence of audible beats (Deutsch, 2013, pp. 24). The *musical* perspective views consonance and dissonance as cultural constructs that change with time according to period and musical style. Before Ars Nova (around 1300) **3rds** and **6ths** were not tolerated as consonances. Between 1450 and 1900 only **3rds**, **5ths** and **6ths** were considered truly consonant, while all other intervals were regarded as dissonant. In many cases the inclusion of dissonant intervals is connected with specific rules for the horizontal continuation such as whether or not a dissonant interval should be resolved to a consonant and so forth. The twelve-tone serialism of the early 20th century in turn completely disregarded the notions of consonance and dissonance all together. Consonance may also be seen from a *psychological* or phenomenological angle and refer to an experienced sense of *”...termination, rest,*

---

[10] *Beats* are an expression of the amplitude modulations occurring between co-sounding tones as a result of partial interference.

absence of tension and a minimal degree of latency and temporality…" (Translated from Danish, M. G. Knakkergaard, Finn, 2003, pp. 1236). This particular perspective is dependent on the former two in that such sensitivities to harmonic tension may be both physically and culturally induced.

The notions of consonance and dissonance may be useful as musical parameters in the context of adaptive VDM in games. As has been suggested in regard to the previously covered aspects of harmony in this section, they may be utilized narratively when coupled to other elements of a game, allowing, for instance, consonant layers of VDM to appear only under specific circumstances. Furthermore, dissonance can - from all of the presented perspectives - be regarded as a form of *harmonic entropy*. This concept will bear some significance in chapter 5.

3.3.5. HARMONY AND VDM

In regards to VDM, I will use the terms *harmony, scale mode, musical key, harmonic anchor, tone cluster* as well as the *consonance-dissonance* dichotomy based on entirely vertically oriented properties - if nothing else is stated, they are regarded as emancipated from their horizontal functionalities of directness and hierarchy - based on the premise that one may experience, for instance, a sense of tonality without necessarily feeling any of the horizontally directed tensions that the terms might otherwise imply. It can not be ruled out, however, that the listener will project directional tensions into vertical harmonic constructs, even when they are presented in a vertically dominated manner. This is to a large extent a question of perception and will be discussed in chapter 5 in the context of evaluating the degree of vertical dominance of a particular music - or angled differently, evaluating the *anticipatory listening position* promoted by the music.

3.4. TIMBRE

The very ambiguous term, *timbre*, is of importance as it is one of the main attributes of the vertical expression. The concept is, however, highly complex and not easily defined even today. Albert Bregman criticizes the American Standards Association's definition for being actually no definition at all. This definition describes timbre as: *"That attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar."* Bregman humorously reformulates that definition to: *"We do not know how to define timbre, but it is not loudness and it is not pitch."* (Bregman, 1994, pp. 92) Nevertheless, timbre is widely accepted as having crucial significance to our ability to identify sound sources - not least through the spectral and dynamic properties of the initial attack portion of sounds.

While subdivisions of the one-dimensional pitch space in terms of tone systems and harmony has largely been based on a distinction between perceived consonance and dissonance the striation of the multidimensional *timbre space* has had to rely on experiments with perceived differences on a multitude of parameters. Some of these are presented in this section.

3.4.1. TIMBRE AND COLOR

A popular striation paradigm for timbre is as a parallel to visual color. The visual color spectrum is physically speaking much more complex than what we perceive as the color circle. The perceptual reduction of color nuances is possible due to *metameric* colors. These color combinations differ in frequency but are perceived as identical. Research in the field of timbre perception seems to not being able to verify the validity of metameric timbre to quite the same degree as metameric color, although some reduction of the spectral continuum is observed in regard to auditory grouping (Bregman, 1994, pp. 122-126). Furthermore, the variety of colors can be understood based on only three dimensions: red-green, blue-yellow and light-dark. Other incompatibilities between color and sound perception are highlighted in (O'Callaghan, 2014) in regard to the relationship between the physicalism of frequencies and our perception of these frequencies as sound and color. I will not, however, go further into these aspects of timbre perception here.

3.4.2. MULTIDIMENSIONAL TIMBRE SPACE AND DESCRIPTORS

Timbre space is multidimensional and the number of dimensions can be seen to vary according to analytical method. Research on models for timbre space is largely built on experiments in which groups of listeners evaluate similarities and dissimilarities of presented sounds. Subsequently the results are mapped in such a way that dissimilar sounds are placed far apart in the multidimensional space of timbre. The sounds used for these experiments range from synthetic sounds, re-synthesized instrument sounds, recorded instruments as well as combinations of recorded instruments (McAdams in Deutsch, 2013, Ch. 2)

To a large extent, the timbre of a sound may be described by its spectral characteristics. Partial distribution and energy balance between partials (accessible through e.g. equalization) are of clear significance. But spectral properties are not the only parameters involved in the research of timbre. The dimensions of timbre space can be represented by a variety of *audio descriptors*. These include parameters such as *spectral centroid, attack time, spectral flux* and *spectral deviation* (McAdams in Deutsch, 2013, pp. 41)

Studies at IRCAM (Institut de Recherche et Coordination Acoustique/Musique) identify 54 audio descriptors categorized into *temporal, spectral, spectrotemporal* and *energetic* sound properties (Peeters, Giordano, Susini, Misdariis, & McAdams, 2011). Temporal descriptors include: Attack, decay, release, temporal centroid, effective duration, as well as the frequency and amplitude in modulation of the energetic envelope. Spectral descriptors include: measures of centroid, spectral spread, skewness, kurtosis (peak sharpness), slope, rolloff, crest factor and jaggedness of the spectral envelope. Spectrotemporal descriptors include: spectral flux. And energetic descriptors include: harmonic energy, noise energy, statistical properties of energy envelope. Additionally, descriptors for harmonicity and noisiness are included.

3.4.3. A PRACTICAL SIMPLIFICATION OF THE NOTION OF TIMBRE

While timbre can be considered a multidimensional phenomenon that should hold all remaining sonorous properties when pitch and loudness have been subtracted, I would argue that simpler approaches to explaining timbre might be useful for practical reasons. It seems evident that most research in the field acknowledges spectral properties to be of special importance to timbre - these spectral properties that differentiate themselves from being separable pitches by perceptually grouping together into an irreducible timbral unity as opposed to the harmony of, for instance, two simultaneously sounding tones, which can be perceptually reduced into its constituting tones. (What Jean Claude Risset calls *spectral fusion* (Kokoras, 2005; Risset, 1991).) The inclusion of *attack times* allows timbre to enter the horizontal domain, and as such I will disregard this otherwise often included parameter of timbre - not least because of its lack of generalizability. Very short attacks may be associated by the listener with a sound's vertical characteristics, but it is arguably difficult to maintain the validity of this parameter when attack times enter the durations of several seconds. After all, it is difficult to see that it should have any timbral implications to change the attack time of, for example, a string sample from 3 to 6 seconds.

One might speculate that on a perceptual level the question of sound as a multidimensional, continuous, smooth-spaced phenomenon may be potentially described according to whichever dimensions we can think of as means of striation. In the context of the view of auditory verticality that has been presented in chapter 2 some of the above parameters will not stand, and in regard to the aim of this thesis to present a frame of understanding on which a system for adaptive VDM can be built, it is clear that a reduction of this potentially infinite problem is necessary.

As McAdams points out, these descriptors show promising potential in regard to sound categorization and automatically performed orchestration tasks (McAdams in Deutsch, 2013, pp. 42). Such categorizations are utilized in, for instance, the *concatenative synthesis* technique developed at IRCAM (Schwarz, Beller, Verbrugghe, & Britton, 2006). From the perspective on auditory verticality argued in the present thesis, however, not all of the identified descriptors above qualify for assimilation into the concept of the *vertical expression*. Some of the parameters that are here attributed to timbre, I will argue, might as well be categorized as properties of texture. (These include *frequency and amplitude in modulation* and *properties of* *the energy envelope*.) And some of them (such as attack, decay, release and duration) have so obvious horizontal implications that I will rule them out all together. A comprehensive study of the compositional exploitation possibilities of the audio descriptors identified for example at IRCAM is beyond the scope of my thesis. A fairly large portion of work lies ahead in claiming control over currently identified as well as future audio descriptors of timbre to be able to implement them in music composition systems in meaningful ways. I will regard such an effort future work and here limit my attention to a few parameters of timbre that seem most likely to be clearly perceivable in a potentially dense soundscape of a computer game. Thus, the parameters of my focus are:

- Spectral envelope (amplitude balance between partials]
- Spectral centroid (spectral position of the averaged energy center)
- Spectral spread (bandwidth of spectrum)
- Spectral flux (evolution of energy in the spectral envelope)
- Spectral deviation (jaggedness of spectral envelope)
- Harmonicity/inharmonicity (presence or absence of a common lower frequency in the proportionate relationship between partials)
- Noisiness (proportionate relationship between noisy and harmonic material)

3.5. TEXTURE

3.5.1. TEXTURE OF MUSIC AND TEXTURE OF SOUND

Another important and equally not easily definable concept when examining musical verticality, the notion of *texture,* has at least a few different meanings as it is used musicologically and amongst music practitioners,which calls for a clarification of its use in the thesis. Based on traditional conceptions, the term can be very generally divided into the *texture of music* and *texture of sound.* I will argue for a broader concept of texture containing both of these conceptions, and include some aspects that are occasionally attributed to timbre, which were ruled out in the previous section.

The notion of texture as a metaphor for qualities of music and sound derives its meaning primarily from the modalities of vision and touch as well as from the physical world of objects where perhaps the most obvious example is the warp and woof of woven textiles. What is traditionally associated with the term *musical* texture are compositional devices such as *monophony*, *homophony* and *polyphony*. Additionally, *heterophony*, *micro-polyphony* and *holophony* (Kokoras, 2005) have been suggested as independent textural types, although they can arguably be assimilated into the first three.

When it comes to the texture of sound, striation attempts include Pierre Schaeffer's sound object studies (Schaeffer, 1966), Xenakis' stochastic methods for the distribution of sound particles[11] or grains (Xenakis, 1992) and Denis Smalley's theory of *spectromorphologies* (Smalley, 1997). But the texture of sound may also be attributed to more traditional means such as instrumental articulation and playing techniques. This implies techniques such as tremolo, vibrato, different pizzicato techniques and flutter tongue as well as many of the expansions to orchestral expressivity offered by among others Krzysztof Penderecki. Sound texture in the context of electronic and digital musical instruments relies on a multitude of parameters including pitch and amplitude modulations as well as those parameters that are associated with the steadily growing variety of DSP effects and other digital sound manipulation techniques.

As is the case with timbre, it is fair to regard the texture of sound as a phenomenon with a potentially infinite resolution of nuance. It is governed by fluctuations in terms of amplitude and spectral properties and might be described very broadly to represent a sound's surface structure in which not least the temporal organization of events are of significance (whereas timbre, as we have delimited it in the above section, is concerned primarily with vertical structure). These events, however, may take a multitude of forms and range from shallow amplitude fluctuations to abrupt timbral changes, and importantly these events are not perceived as separate entities but precisely as a cohesive surface structure, grouped together by the gestalt principle of temporal proximity as described by Bregman (Bregman, 1994, 2005) and discussed in chapter 2. The texture of sound is a property of the micro time scale. Musical texture on the other hand is taking place at the meso time scale but is concerned with principally the same horizontal and vertical organization of events - only here, these events are easier to separate and any underlying structural patterns are more immediately accessible to the listener because they evolve at a slower pace.

3.5.2. TEXTURE ACCORDING TO TIME SCALE

I will regard *texture* as broadly applicable to all sound, be it music or otherwise, and graduate - or striate if you will - this broader conception of texture into subdivisions according to Curtis Roads' time scales. *Meso-texture* in this respect refers to constructs within the meso time scale and cover relations and juxtapositions of musical ideas, monophony, homophony, polyphony, chord progressions, the variation, imitation and development of motives, themes and melodic sequences etc. *Sound object texture* refers to the dynamic, harmonic and spectral morphology within the timeframe of single sustained sound objects or tones. *Micro-texture* is

[11] The idea of sound quanta or sound atoms was investigated in the 1940s by physicist Dennis Gabor (Gabor, 1947), but can be traced back to the Greek atomistic philosophies of Leucippus and Democritus in the fifth century BC.

concerned with the arrangement of sonic material at the micro time scale including the quality and distribution of sound particles, grains as well as spectral and dynamic fluctuations. Furthermore, phenomena such as pitch and timbre can be understood to arise as an expression of micro-textural organization. Micro-texture falls within the timeframe of what we have called the *vertical extension* and is thus, as the only one of the mentioned texture types, a constituent of the *vertical expression.*

3.5.3. SPECTROMORPHOLOGIES AND TEXTURE

Pierre Schaeffer's elaborate system for the categorization of sound objects is much quoted and receives special status among scholars and composers because of its role as the origin of *music concrète* - from which many experiments with electroacoustic music have since emanated and taken inspiration. The descriptive and non-technical nature of Schaeffer's terminology, however, has made its practical applicability less than optimal and being conceived in the cradle of electroacoustic music, more recent endeavors for the striation of verticality in terms of texture and timbre exist. One of these is scholar and electroacoustic composer, Denis Smalley's theory of *spectromorphologies* - a term similar in meaning to Schaeffer's sound object, and which also owes much of its conception to precisely Pierre Schaffer's initial studies. Smalley coins the term to account for musical structures, which are, as he puts it:

> "…more interested in spectral qualities than actual notes, more concerned with varieties of motion and flexible fluctuations in time rather than metrical time, more concerned to account for sounds whose sources and causes are relatively mysterious or ambiguous rather than blatantly obvious." (Smalley, 1997, pp. 109)

The spectromorphological approach entails both vertical and horizontal shaping ideas of electro-acoustic music. Although conceived primarily in relation to electroacoustic music, many of the concepts involved are easily applicable to other musical styles and are generally very well suited to describing the structure of VDM as an umbrella concept of musics that do not necessarily owe a debt to traditional score based representation and therefore to the analytical methods traditionally associated with these. In the following sections are presented some of those aspects of Smalley's theory that can lend themselves as structuring paradigms primarily on the vertical dimension.

3.5.3.1 Internal and external note-view

Smalley divides notes of fixed pitch into two note-views: *internal* and *external spectral focus* (Smalley, 1997, pp. 119). The *external note*-view covers the traditional musical note where the pitch is the most important feature. For *internal note* focus, the interior components of the instrumental note - its spectral content of partials and their relative energy distribution, are examined. This focus hereby ”moves inside the note”, as he puts it. Smalley mentions close mic’ed recordings of instruments as a way of allowing such sounds to be manifest.

Apart from the occupation of spectral space through notes, Smalley speaks of two kinds of noise in regard to his work as an electroacoustic composer (Smalley, 1997, pp. 120). *Granular noise* is described as a texture constituted by impulses or grit. Examples of associations coupled to granular noise include the sea, wind, rubbed and scraped materials. The other noise category he terms *saturate noise* described as a saturation of a portion of spectral space to the extend to which no pitch perception is possible - neither intervalic nor relative.

In the area between note and noise sits the phenomenon of *inharmonicity* defined as a quality to sounds (such as bell sounds) for which overtones are not multiples of a common lower frequency as is the case with *harmonic* instruments like strings, flutes and human vocals. Importantly, notes can become saturate noise through increased inharmonicity and vice versa, making inharmonic sounds a valuable middle ground between noise and note.

Smalley goes on to further describe the possibilities of structures in spectral space with four qualitative dichotomies of spectral occupancy (Smalley, 1997, pp. 121): *Emptiness - plenitude* covering the degree or amount of occupancy in the spectrum. *Diffuseness - concentration* describes whether sound is dispersed or concentrated in regions within the spectral space. *Streams - interstices* refer to the layering of spectral space into narrow or broad streams separated by intervening spaces. *Overlap - crossover* describes how streams or spectromorphologies encroach on each other’s spectral space, or move around or across each other to another region.

3.5.3.2 Texture motion

Smalley identifies a number of qualifiers in regard to the motion of the internal texture of individual spectromophologies. Firstly, textures may move in a sustained *continuous* or a fragmented *discontinuous* manner (and they may do so by either *streaming*, *flocking*, *convoluting* or *turbulent* type texture motion, but I shall leave this aside for the moment as these concepts are associated with motion at the sound object time scale and thus fall outside the present chapter’s focus on the vertical expression.) Micro-textural motion may be *discontinuous* due to *iterative* or *granular* textural qualities - the space between which ranges from periodic repetition to the erratic (which I have earlier referred to as *cyclic* and *non-cyclic stasis*). *Continuous* texture motion spans from sustained to granular - with granular texture acting as a common middle ground, which, depending on the temporal density of grains, may be perceived as either continuous or discontinuous. Except for purely sustained textures, texture motion, generally speaking, may in turn be periodic or aperiodic, it may *accelerate*, *decelerate* or be in a state of *flux*, and it may, in Smalley’s conception, also under some circumstances imply higher level groupings into patterns. Whether or not these patterns are perceived as patterns is not discussed but most likely what Smalley refers to here are patterns that reside within the same perceptual window (as covered in chapter 2) and thus do not form separate gestalts as this strictly speaking would imply the formation of a new spectromorphology and no-longer be an artifact of internal texture motion.

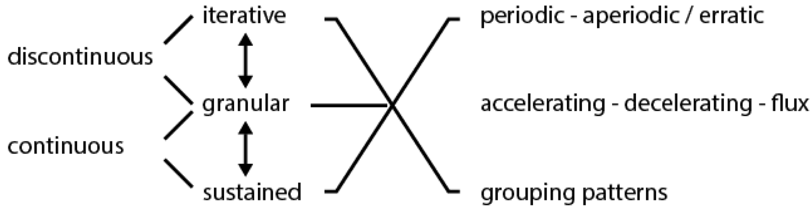


Figure 5 - Texture motion.

Smalley's spectromorphological approach is not alone as a viable striation paradigm for the texture of sound. It, however, manages to encompass several other approaches. For instance, the range from iterative over granular to sustained texture motion is often elsewhere referred to as the degree of texture *density* (e.g. Xenakis, 1992), or as a range from *smoothness* to *roughness* (Roads, 2004, pp. 341). The smooth and rough in Roads' terminology is an analog to the harmonic concept of *sensory dissonance* in which *beats* may produce such rough tensions. Roughness is a term used by many both scholars and practitioners (e.g. Iverson, 2010, pp. 66; Truax, 1982, pp. 76; Vaggione, 1984, pp. 49). It seems, however, less precise in its metaphorical approach than Smalley's more neutral terminology, and could arguably be understood as slightly charged - giving potential connotations of unpleasantness, which might or might not be illustrative of a particular granular texture.

Lastly, in relation to Smalley's spectromorphologies, William W. Gaver deserves mentioning. He presents an approach to understanding sound in the natural environment by categorizing them according to attributes of the sound-producing event[12] that cause them (Gaver, 1993). He speaks of three primary categories: Vibrating solids, aerodynamic events and liquid sounds. I will not go further into Gaver's ideas here, however, due to their emphasis on sound source and the concrete events that cause them. What I am interested in, rather, is concrete parameters that can be changed in real time as a means of evolving vertical expressions in computer-generated VDM.

---

[12] Gaver does not refer to these as spectromorphologies.

3.6. AURAL SPACE AND AUDITORY VERTICALITY

As a last subject in this chapter, I must briefly address the question of auditory spatiality. Arguably, acoustic properties in the form of reverberation and delay may be understood as qualities of auditory verticality - qualities of the vertical expression. The sense of space can be argued to be a vertical experience even though spatial effects such as echo and reverb have a clear temporal base. Reverb tails unfold over time, which is structurally an expression of horizontality, but the sense of space that arises in the listener can be seen perceptually as a vertical phenomenon. Long delays can be different in this respect if heard as a regular decay of clearly separable echoes that exceed the micro time scale. Such separable and obvious repetitions thus become a clear expression of auditory *horizontality*.

Barry Blesser mentions *social*, *navigational*, *aesthetic*, and *musical* spatiality in his phenomenological perspective on space perception - which he refers to as *aural space.* He writes:

> “The physical acoustics of a musical space merge with sound sources to create a unified aural experience. Space then becomes an extension of the musical or vocal art form performed within it.” (Blesser & Salter, 2009, pp. 11)

This notion of a ”unified aural experience” seems to support the conception of perceived spatial properties of sound as co-constituents of auditory verticality.

Although there are an immense number of different reverb effects on the market, typical spatial parameters that can be found on most DSP reverbs include:

- **Reverb type:** This parameter typically includes algorithms for different kinds of rooms such as hall, room, chamber and church as well as different reverb technologies including plate reverb, spring reverb, reversed and gated reverbs
- **Room size:** Changes the size of the virtual acoustic space.
- **Pre-delay:** Determines the duration of an initial delay before which sounds are left unaffected.
- **Decay time:** This parameter sets the duration of the reverb tails decay.
- **Diffusion:** The diffusion parameter can be explained as the complexity of the reverberation in terms of how perceivable are the individual echoes that make up the reverb tail. High settings create a smooth an saturate sounding reverb tail, while lower settings may render the reverb tail more ”transparent”.

Physical and perceptual properties of acoustics is a vast field - a thorough investigation of which is beyond the scope of this thesis. I will confine myself here to merely point to the existence of this spatial aspect of auditory verticality and its perceptual, and thus also *narrative* potential, that must not be overlooked in the context of game music. Narrative considerations on spatiality are taken up in chapter 6 - not least in regard to reverb tails with a perceptually infinite duration.

I will not, however, include acoustic properties into my definition of the *vertical expression* as such. I omit it because its function seems to differ in comparison to harmony, timbre and micro texture. It may be seen to act more as a medium for these vertically expressive components of the vertical expression to ”sound in” than as a part of the sound source itself. Rather, I shall regard it as an intricately connected aspect of auditory verticality, which is ever-present as a companion to any sounding manifestation of vertical expressions.

3.7. CONCLUSION

As mentioned in the introduction, this chapter has two primary purposes: To identify vertical parameters that can be used to generate VDM in computer games; and to provide a definition for the notion of vertical expression introduced in the previous chapter.

Provided with the above non-comprehensive review of vertical striation paradigms it is possible to identify a number of parameters, which may be set to control the composition of vertical expressions. These parameters, which will be subject to additions as more perspectives are presented in the thesis, can for now be categorized into parameters of *tone systems, harmony, timbre, micro-texture* and *aural space.*

Firstly, compositional decisions can be made in regard to *tone system*. Parameters within this group include (but are by no means limited to) the Pythagorean and pure tunings as well as various forms of temperaments including mean-tone temperament and 12 tone equal-temperament. Additionally, the octave may be divided into micro-tonal temperaments of any number of subdivisions such as quarter- and eighth tones. Finally, it is possible to respect the vertical auditory dimension as a smooth space and regard it as an open spectral continuum or divide it based on the overtone series as is the case in spectral music. (Spectral music will play a significant part in the next chapter.)

Secondly, pitch relations within a chosen tone system may be subject to certain schemes of organization or *harmony*. In 12 tone equal-temperament these include traditional phenomena such as *scales* and *modes, musical key,* possibilities of *polytonality*, *harmonic base intervals* (such as 2^{nd}, 3^{rd} or 4^{th} harmony) as well as *tone cluster*s, which may be harmonically specified in terms of their *register* and *register range*. In regards to VDM, the notion of musical key can be seen to have a variant that I have referred to as *harmonic anchoring*. Harmonic anchoring may take the form of either a single tone, a chord of tones or a cluster of closely spaced tones. Moreover, the concepts of *consonance* and *dissonance* in both the sensory and musical conceptions can be put to compositional use in the formation of vertical expressions. Interestingly, harmonic properties of verticality may give rise to the formation of one or multiple auditory streams. The *number of auditory streams* can in itself be regarded as a vertical parameter.

Timbre, as I have treated the term in the present chapter and will refer to it in this thesis, confines itself to primarily spectral properties of sound and includes parameters associated with the *spectral envelope* of a sound such as *spectral centroid, spectral spread, spectral deviation* and *spectral flux.* Some of these timbral qualities can be accessed through normal equalization, while others require more radical methods. Also the phenomena of *harmonicity, inharmonicity* and *noisiness* go under this category. Importantly, it is clear that higher level concepts such as *instrumentation* and *orchestration* play a significant role in the formation of timbre. This refers to acoustic instruments as well as other sound generating techniques such as varies forms of synthesis, sampling and noise generators.

The last category of auditory verticality presented in the above deals with the concept of *texture*. I propose to divide the notion of texture based on time scale. As such, mainly *micro-texture* is of interest in terms of the vertical expression. Parameters of micro-texture may be found in traditional and modern instrumental *articulations* and *playing techniques* in the form of tremolo, vibrato and flutter tongue to name a few. (The next chapter provides examples of expansions made to the traditional orchestral expressions by composers like Penderecki and Ligeti.) This category also contains *pitch* and *amplitude modulations* as well as various *DSP effects* with micro-textural implications. Striations of the infinite vertical space finds a useful terminology in Denis Smalley's *spectromorphological* approach. Here the vertically oriented aspects of texture, as presented in this and the previous chapter, are described by the way in which the texture is internally in motion. Smalley speaks of *continuous* and *discontinuous* texture motion, which may span from *iterative* to *granular* and further from granular to *sustained* motion. Additionally, this texture motion may have *periodic* or *aperiodic* characteristics, it may *accelerate*, *decelerate* or be in a state of *flux.* Furthermore, spectromorphologies can exhibit *saturate* and *granular noise*.

On a more general note, the occupation of spectral space can be regarded via an *internal* or *external note-view* and happen to various degrees within opposites of *emptiness-plenitude*, *diffuseness-concentration*, *streams-interstices* and *overlap-crossover.*

The notion of *aural space* was discussed as an intricately connected aspect of auditory verticality - an ever-present companion to sounding manifestations of vertical expressions. Although I do not include spatial properties into the definition of the vertical expression I have listed a number of basic parameters for the control of this auditory vertical aspect. These are: *Reverb type, room size, pre delay, decay time* and *diffusion.*

The vertical parameters identified in this chapter can be found - together with vertical parameters identified in the remaining chapters - in Appendix F.

It is with a basis in these parameters and within the framework put forth in this chapter in general that the term, *vertical expression*, can be understood as: *an emergent hybrid concept of pitch, harmony, timbre and micro-texture*.

The vertical dimension is a fundamental aspect of musical expression. After all, any musical artifact relies on some means of manifestation and this, whether it is an instrument, a voice or a teapot is inextricably linked to a certain potential for vertical expressiveness. As discussed briefly in chapter 2, the degree of focus on this aspect of music, however, has changed gradually over time to take an increasingly stronger role. In the next chapter we will look into a few examples of how music that is dominated by this dimension has taken form by the hands of different composers, some of whom were occupied with VDM in the mid 1960s, while others represent more recent endeavors.

CHAPTER 4. THE STRUCTURE OF VDM

4.1. INTRODUCTION

Following up on the previous chapter's look into the inner structure of *vertical expression*, the present chapter serves to further elaborate on the structural part of a structural, aesthetic and functional foundation that may allow for the development of a generative system for implementing adaptive VDM in interactive contexts. It does so by presenting common underlying structural properties of VDM and in turn argues for VDM to be used as an umbrella term to encompass a number of musical styles based on such properties. I will argue that the historical development of verticality in composition has seen a gradual emancipation of musical structure from strictly rule-based occupation of pitch space to a free access to spectral space as the grounds for composition.

In this chapter I present common structural tendencies associated with *three general characteristics* for facilitating vertical dominance as these have been expressed in existing VDM: the *avoidance of horizontally separable gestalts*, *avoidance of perceivable horizontal regularities* and *a weakening of the perceived bond between spectromorphologies*.

A look at the concept of *entropy* is presented. I argue for the applicability of the notion of *musical entropy* as a measure of the perceptual degree of musical disorder and examine on this basis *harmonic, melodic, rhythmic* and *timbral entropy*. From here I go on to argue for applying the Deleuzo-Guattarian notion of the *rhizome* as an analogy to the musical structure of VDM.

The three general characteristics of VDM and the concept of musical entropy are demonstrated along with a set of techniques for the *expansion of vertical expressivity* by examining compositional techniques from a selection of composers (Geörgy Ligeti, Krzysztof Penderecki, Iannis Xenakis, Gerard Grisey and Tristan Murail), who have in some parts of their work embarked on VDM composition and who can be considered pioneers of the style. A comprehensive study of VDM styles, methods and composers is however not presented. The focus is, rather, on identifying general underlying principles of VDM structure, which have found a myriad of different specific manifestation methods of which only a limited group is included here. Special emphasis in this respect is put on Geörgy Ligeti, who in some of his works of the early 1960s outlines, in an exceptionally clear manner, characteristics that are common to also other styles of VDM. Ligeti's works, *"Atmosphêres"* (1961) and *"Lux Aeterna"* (1966), furthermore play a significant role in the non-diegetic music of one of space-oriented science fiction's classic movies, Stanley Kubrick's *"2001: A Space Odyssey"* (Kubrick, 1968), making these works well suited for the overall subject of this thesis.

The notion of musical entropy is used as an analytical parameter in a summary of analysis, which examines the parts of Ligeti's works,*"Atmosphêres"* and *"Lux Aeterna"* that were used in *"2001: A Space Odyssey"*. Aspects of Ligeti's micro-polyphonic technique are examined. Penderecki's experiments in playing techniques and Xenakis interest in probability as a basis for music composition as well as the relationship between continuous and discontinuous textures by way of, for instance, granular synthesis are discussed.

The chapter goes on to arguing that the use of Penderecki and Ligeti on film associates their VDM styles with filmic space-oriented science fiction, horror and mystery, giving a hint to how this music may impact upon the gaming experience. This is, however, a subject that will be taken up in chapter 6 and will not be elaborated at any length here.

A range of techniques, concepts and methods of spectral music is presented in order to draw out additional vertical parameters for an initial set of generative compositional principles. These include not least methods associated with *orchestral synthesis* and the idea of compositional *reservoirs*. A key aspect of VDM, as exemplified by the composers mentioned in this chapter, is the relationship between this music's perceptual *un-specificity* and the immensely nuanced notational *specificity* that govern it. This is discussed in the concluding part of the chapter.

4.2. THE EMANCIPATION OF VERTICALITY

Historically speaking the focus on the vertical musical axis - the sound in itself - can be seen to have gradually grown and deepened with time. Examples of the vertical expression beginning to take a more central role are found in the first half of the 20th century with composers such as Edgard Varèse, Olivier Messian, Henry Cowell and John Cage. This gradual movement would by the second half of the 20th century lead to necessary expansions of the vertically expressive capabilities of the orchestra - not least through the inclusion of new playing techniques. Such techniques include for example col legno, sul tasto and sul ponticello and the use of harmonics for string instruments; flutter tongue, blowing with no pitch and producing sounds with instrument mechanics; prepared piano and strumming and picking directly on the piano's strings; new strokes as well as the rubbing and scraping of percussive instruments; vocal whispering, yelling and speaking. With the invention of the tape machine, music was liberated from the vertically expressive limitations of the orchestra by allowing composition with recorded and generate sound objects that could be stored on tape - paving the way for the use of tape loops as a means of composition for the *musique concrète* by the composers Pierre Schaeffer and Pierre Henry in France, and providing a medium for the *elektronische music* of, for instance, Karlheinz Stockhausen in Germany to be presented on.

However, the emancipation of musical verticality is not only concerned with the more noisy sonorous experiments of such composers but takes root (at least in the western tradition) in the evolution of polyphony. This development can be traced through the introduction of parallel 4th and 5th harmony in Gregorian chant music, paradigms for the treatment of dissonant intervals as in Palestrina's polyphonic technique, the development of the functional harmonic system through the baroque and classical eras, and its gradual chromatic liberation in the romantic era, increasingly free conception of harmony with impressionist composers like Debussy, and a total disintegration of tonal hierarchy with serialism. In the 1950s, 1960s and 1970s this development saw a breaking with the 12-tone subdivision of the octave into micro-tonal temperaments with sound-mass music, sonorism and spectral music. During the last 50 years this liberation has only accelerated with the possibilities offered by new sound technologies. Overall, it is fair to say that the evolution of verticality exhibits a gradual emancipation of tightly rule-based striations of pitch-space into an increasingly free utilization of spectral-space for music composition. This, of course, also has implications for popular music as an increasing focus on production, which has its perhaps most salient expression in some examples of IDM (Intelligent Dance Music) and other forms of contemporary electronic music.

But vertical dominance is not the same as an emancipation of verticality that allows new sounds to be accepted as musical. It is entirely possible to compose highly horizontally dominated music with sounds that are unorthodox in a musical context. Styles associated with such musical directions as sonorism, sound-mass music and spectral music as well as some minimal music and several other styles both western and otherwise may, according to specific music-structural characteristics, be gathered under the stylistic umbrella term, *vertically dominated music*.

4.3. COMMON STRUCTURAL CHARACTERISTICS OF VDM

4.3.1. THREE GENERAL CHARACTERISTICS OF VDM

The common structural characteristics of VDM can be found in the compositional means that facilitate the music's vertical dominance. On a general level, I will argue that vertical dominance may be achieved due to three general characteristics:

- avoidance of horizontally separable gestalts
- avoidance of perceivable horizontal regularities
- weakening of the perceived bond between spectromorphologies

These are, as I write, characteristics, and not rules. And this is an important point as it is not necessary for all of them to be met entirely for vertical domination to arise, just as a piece of music may have an overall characteristic of being vertically dominated while also containing passages of horizontal domination. To avoid an overly categorical conception of vertical dominance it can be explained as an expression of the angle of the musical ”diagonal” leaning to various extents towards the vertical axis.

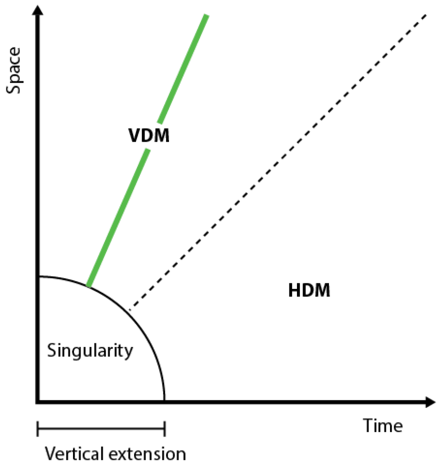


Figure 6 - Musical “diagonal” (in green) leaning towards VDM.

The term VDM is not meant to refer to the ideal state of a music that is purely[13] vertical. Furthermore, it is often not enough for only one characteristic to be met. In most cases vertical dominance occurs as a result of combinations of these musical characteristics that also often overlap.

Firstly, the avoidance of horizontally separable gestalts has been achieved in existing VDM by mainly two methods: 1) by obscuring the onset and termination of the spectromorphologies that form the meso-texture and 2) by achieving a mass-merging of many individual events. Secondly, avoiding perceivable horizontal regularities is on a structural level manifest broadly in VDM as 1) an absence of repetition, rhythmic pulse and accentuation of meter, and 2) as an absence of patterns in the form of themes, motifs and rhythmic figures. Thirdly, a weakening of the perceived bond between spectromorphologies may be achieved structurally by 1)

[13] The only purely vertical musical entity, based on how auditory verticality has been defined in the previous chapters, is the *vertical expression* itself residing within the duration of the *vertical extension.* An extreme case of vertical extension exists in the aforementioned work by La Monte Young,”Compositions 1960 #7”.

an absence of causal direction between spectromorphologies, and 2) by letting the duration of spectromorphologies or the pauses between them exceed the period at which perceived event cohesion is likely to occur. (As discussed in chapter 2, this duration ranges from approximately 2 to 5 seconds as implied by Daniel N. Stern conception of a *present moment*, Wittmann's *experienced moment*.)

- **Avoidance of horizontally separable gestalts**
    - By obscuring of onset and termination
    - By mass-merging of many individual events
- **Avoidance of perceivable horizontal regularities**
    - Due to an absence of repetition, rhythmic pulse and accentuation of meter
    - Due to an absence of patterns like themes, motifs and rhythmic patterns
- **Weakening of the bond between spectromorphologies**
    - By an absence of causal direction
    - By exceeding the period of event cohesion

4.3.2. EXPANSION OF VERTICAL EXPRESSIVITY

Apart from the above mentioned three general characteristics that seek to *avoid* certain musical traits, VDM is characterized by an expressive emphasis on the *vertical expression* rather than traditional horizontally dominated musical effects, and approaches this emphasis through extending the vertically oriented musical means. These expansions, of which examples will be covered later in the chapter, include:

- New orchestral playing techniques
- Emphasis on mass-effect, rather than individual pitches and rhythms
    - Via clouds of a large number of events with high horizontal density
    - Via tone clusters of high vertical density
- New musical texture conceptions
    - Continuous and discontinuous textures
    - Micro-polyphony
- Experiments in timbre
- New orchestration techniques
    - Orchestral synthesis

  - Micro-tonalities
  - Very large polyphony through total divisi orchestration

- New technologies
- New horizontal development schemes

  - Gradual development of tone clusters
  - Gradual development of vertical parameters
  - Irreversibly changing static musical time

- A focus on the perceptually *unspecific* (which can be structurally very specific through meticulous notation)

With only few exceptions all of the above mentioned characteristics, expressive expansions and characteristics are present in Ligeti's *"Atmosphêres"* (1961) and *"Lux Aeterna"* (1966) which I will take a closer look at later in the chapter. But first I must clarify the use of the term *musical entropy*, as it is of some significance in the present chapter as well as the next one.

4.4. MUSICAL ENTROPY

The notion of *entropy* is expressed in the second law of thermodynamics as a degree of disorder and uncertainty in a system. High entropy describes a high degree of disorder, while low entropy describes the opposite, a high degree of order. With time, however, the concept has been used in art theory and information theory to describe other subjects than thermodynamics. The notion of entropy is distinguished by a certain unequivocal and universal approach to the oppositions of order/disorder, organization/disorganization, which makes it suitable as an overarching and unifying expression that, moreover, has the ability to act as a common thread across gaps between structure, experienced structure and universal principles in the literal sense, which serve the subject of this thesis well. Although I also use the notions *order* and *organization*, these are less suited as overall descriptors. The word, *order*, creates potentially unwanted associations to a division into hierarchical and delimited levels, while *organization* seems to hint at intention, which may be irrelevant and at worst misleading. When I use entropy here in a musical context, it will be as a measure of the perceived level of sonorous or musical disorder or disorganization at several clearly defined levels.

Transferred to the realm of sound, the roar of the ocean waves or the wind in the leaves, both of which are characterized by randomness, disorder and disorganization, can be considered as examples of extremely high sonorous entropy. On the other hand, the interval of an octave or a beat with steady pulse would stand as examples of very low sonorous entropy. If we, for now, stick to Varèse' definition of music as *organized sound* it is problematic from a structuralist perspective to imagine music with a high degree of entropy. Music is in that case precisely *organized* and as such inherently characterized by low entropy. From a phenomenological point of view, however, where it is the *experienced* disorganization that is in question, things look quite different. A good example of this can be found in serialism, which is often criticized for an inconsistency between its compositional structure and a perceived lack of structure. The experience of the world implies a natural involvement of the subject. Some will experience a piece of music as chaotic or disorganized while others hear the same music as orderly and organized. Therefore, the phenomenological approach to entropy that I argue for here is tied to the relative probability of experienced order or disorder in a field of tension between overt and covert musical structure, which in turn depends on the listener's music-cultural background. It is probable that the repetitive harmonic progression of a 12 bar blues will be experienced as more ordered than a complex and unpredictable musical development of a piece of serial music, but it is likewise probable that Indian raga improvisation is experienced as more organized by someone from India than by a Westerner.

The roar of the ocean and the octave interval each belong to their own extreme, but many sonorous and musical factors can be categorized according to the degree of perceived entropy they facilitate. Let me mention here a few examples of this phenomenological approach to musical entropy with reference to four basic music-analytical parameters; melody, harmony, rhythm and timbre.

High *melodic entropy* can be expressed by a low level of perceived melodic cohesion. A tone row that does not clearly exhibit a hierarchy of pitches, thematic work, imitation, repetition or by other means takes audible point of departure in conditions from earlier in the piece to form its further melodic development, can be described as having a high melodic entropy, a high degree of perceived melodic disorder.

A good example of low *harmonic entropy* is the tradition of tonal harmony wherein any harmonic structure, ideally, is set in the world by virtue of its function in relation to a harmonic home, the tonic. Also the degree of dissonance in a given harmony, can have high or low entropy determined by the regularity or order in the harmony's overtones. An ***F-major*** triad has low harmonic entropy in regard to overtone coherency compared to a chromatic tone cluster spanning from ***f*** to ***c*** - or for that matter a vertical expression governed by spectral compression or stretching of overtone distribution. These are clearly perceived as consonance and dissonance respectively taking into account the implications of dissonance discussed in chapter 3.

High *rhythmic entropy* can be exemplified by shifting time signatures, absence of clear pulse and tempo indication as well as irregular timings in relation to meter, or an absence of meter altogether.

An example of high or low *timbral entropy* could be high or low level of timbral uniformity or consistency throughout registers. If a talented violinist plays all notes within the instruments register in a homogeneous timbral manner despite string and position shifts, this can be regarded as an expression of low timbral entropy. In such a case, the formant pattern is ideally reserved throughout, in other words exhibiting a consistency of organization in regard to relative overtone amplitude. On the other hand, one can imagine an assembly of several different recordings of different ensembles in differing rooms as an example of a sonorous expression, characterized by high timbral entropy, by uneven timbral properties, which do not in any way adhere to the same timbral order or organizational principal. The mastering process in music production deals, among other things, with the shaping of timbral material of relatively high entropy towards lower entropy through equalization and multi-band compression according to certain aesthetic ideals.

Audio production and sound design in general basically deal with the order or organization of sound in terms of both spectral balance by equalization, dynamics by compressing or expanding, stereo image by panning and psychoacoustic stereo expansion, the sound image depth by a mixture of equalization, panning and addition of effects like delay and reverb, why the notion of entropy is particularly suitable also in this context.

Several other examples could have been mentioned above. I would generally argue that the notion of entropy lends itself to any musical context as a relative measure of the degree of perceived disorder and disorganization - with the reservation that cultural aspects that have implications for the listener's musical background play an in some cases major role.

4.5. VDM AS RHIZOMATIC MUSICAL STRUCTURE

Another useful analogy to VDM's musical structure is the notion of the *rhizome.* In connection with their elaborations on smooth and striated space-time in ”A Thousand Plateaus”, Deleuze and Guattari refer to the rhizome as a metaphor for their ”new image of thought”. The term originates in the plant world where it is associated with some subterranean root structures such as potatoes and ginger. I will argue that the rhizome is a viable metaphor also for the structure of VDM.

Edward Campbell writes about the Deleuzo-Guattarian application of the term:

> ”The new image of thought is also conceptualized with the vegetal model of the rhizome, and stands in opposition to the traditional image of thought, which is defined as arborescent. In contrast with the hierarchical structured branches found within tree systems, a Deleuzo-Guattarian rhizome has lines which allow the connection of any of its points with any other, and where arborescent systems have ”hierarchical modes of communication and established paths, the rhizome is an a-centred, non-hierarchical, non-signifying system.” (Campbell, 2010, pp. 143)

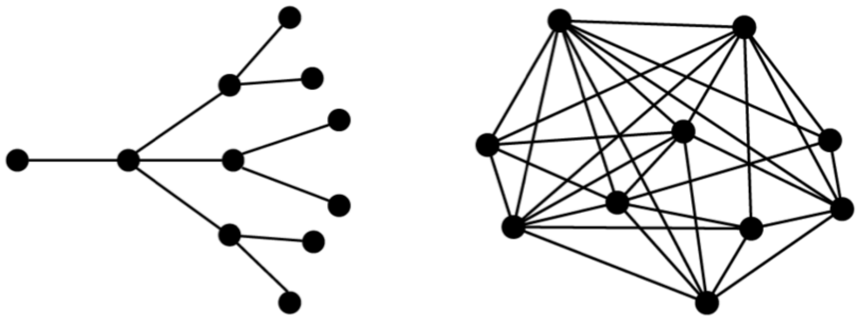
Rhizomatic structures are, as opposed to arborescent structures (tree- structures) not governed by hierarchy - such as the hierarchy of stems and branches. In a rhizome everything is connected to everything and there is no pivot. Deleuze and Guattari write:

> ”…any point of a rhizome can be connected to anything other, and must be. This is very different from the tree or root, which plots a point, fixes an order.” (Deleuze et al., 1987, pp. 7)

And further:

> ”…a rhizome is not amenable to any structural or generative model. It is a stranger to any idea of genetic axis or deep structure.[…] It is our view that genetic axis and profound structure are above all infinitely reproducible principles of *tracing*. All of tree logic is a logic of tracing and reproduction.[…] The rhizome is altogether different, a *map not a tracing.”* (Deleuze et al., 1987, pp. 12)

The rhizome structure, in making a *map* rather than *tracing* a logic, may in this respect, I argue, be seen as an ultimate expression of *verticality*. By not *plotting a point* or *fixing an order* and when all points must connect to all other points, rhizomatic structure is intact in its infinitude and irreducible into constituting substructures.



Figur 7 - Simplistic illustrations of arborescent and rhizomatic structure.

In the context of VDM structure, the characteristics of the rhizome highly resemble the characteristics identified earlier in this chapter such as the concept of mass-merging and the obscuring of onset and termination associated with the avoidance of horizontally separable gestalts. A tree structure, on the other hand, being subject to a generative model (i.e. a stem splits into branches that again split into branches and so on - a simple generative concept) is seen to reproduce itself. This image goes well in hand with the concepts of causality, repetition and horizontal regularities associated with HDM, where musical material refers back to itself in processes of self-similarity (such as imitation, transposition and thematic development). Furthermore, tonal HDM exhibits a tonal hierarchy with a clear tonal center analog to that of an arborescent pivot or stem. The rhizome, by being *not amenable* to such *deep structure,* resembles the behavior of VDM with this music's high degree of musical entropy and tendencies to avoid such hierarchic order and causal logic.

While an arborescent structure can be seen as an expression of directionality in time, and thus an expression of hierarchy and horizontality, the rhizome is timeless and neutral in this respect. Rhizome structure is also an excellent metaphor for highly *non-linear* game designs such as "sandbox" games where the narrative is not controlled along a tree structure, but where players are allowed to roam freely within the functionality of the game. This is an aspect of the game media that I will discuss in some detail in chapter 7.

For now, I will continue my investigations on VDM structure by examining two works by Hungarian composer Geörgy Ligeti in which the general characteristics of VDM highlighted earlier, play a very significant role.

4.6. AN INSIGHT INTO GEÖRGY LIGETI'S "ATMOSPHÊRES" AND "LUX AETERNA"

For demonstration purposes a short summary of analyses of the sections in the two works by Ligeti that were used in the science fiction film *"2001: A Space Odyssey"* (Kubrick, 1968) is presented in the following pages. In this film, and a number of others, VDM has been used to accompany scenes with a focus on the narrative setting of outer space, infinity, mystery, paranormal activity and a presence of "something" that exceeds the boundaries of the senses. The analyses themselves, which are based on an adaptation of David Cope's vector analysis method (Cope, 1993), are not presented here. For detailed analyses of the passages of *"Atmosphêres"* (Ligeti, 1961) and *"Lux Aeterna"* (G. r. Ligeti, 1966) that are used in the film, see my masters thesis (Kristensen, 2011). The summaries that are included here highlight aspects of the passages that are concerned with the following areas:

- Filmic context
- Overview, structure and meso-texture
- Orchestration techniques
- Basic techniques
- Vertical models
- Horizontal models

Finally, the most significant musical traits of the passages are presented in a bullet point overview.

4.6.1. ”ATMOSPHÊRES”, G. LIGETI, (1961)

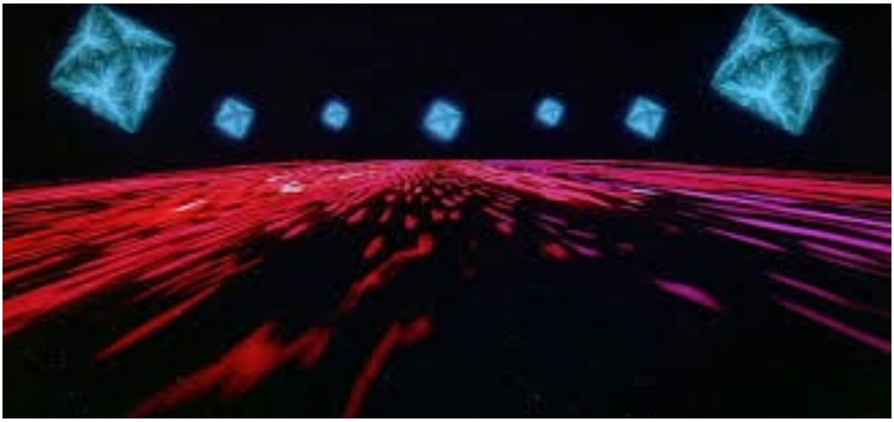


Figure 8 - Screenshot from the chapter ”Jupiter and Beyond the Infinite” in ”2001: A Space Odyssey”

4.6.1.1 Filmic context

”Atmosphêres” is present at three occasions in the film. The first 2-3 minutes of the work (until letter ”D” in the score) accompanies the beginning of the film, that on the image side, consists of a black screen. The second part of the movie (after a pause which splits it on the middle) starts in the same way, also accompanied by the first minutes of *”Atmosphêres”*. The third time *”Atmosphêres”* appears it is in nearly its full length in the chapter of the film titled ”Jupiter and Beyond the Infinite”, which presents a visual journey through an undefined seemingly infinite space full of colorful lights and shapes. The scene is the culmination of the entire film's focus on the mysterious so-called ”Monolith” - a central theme in the movie in the form of a symmetrical, rectangular, 3-dimensional shape which meaning is unknown. The music's function is as a leitmotif or symbol for the unknown, infinite and mysterious ”something” which is the focal point of the film.

4.6.1.2 Overview, structure and meso-texture

The overall form of the piece is best described as a continuous and irreversibly changing flow of textures, which general dynamic shape draws an arch from a subtle *pianissimo* onset *crescendoing* towards a more powerful middle section, which in turn decreases in intensity towards a long *morendo* at the end. The work is generally characterized by an alternation between different cluster-based structures that internally are harmoniously static but are rich in variation in terms of timbre, dynamics and micro-texture and exhibit an unconventionally large dynamic nuance from *pppp* to *ffff*. It is often impossible to distinguish individual voices from the whole they are part of, and the often fully chromatic tone clusters appear as a mass of sound that only rarely gives rise to the experience of chords in the traditional sense and - perhaps with the exception of a complicated web of motifs and tremolos on the interval of a minor third (ex. M.23-29) - never to melodic or rhythmic motifs or harmonic progressions as such. Rhythm is generally characterized by an absence of musical pulse and an avoidance of any emphasis on the meter and obvious tone onsets, which is also underscored by direct instructions in the score such as *”Dolcissimo, imperceptible attack"* and in many places, carefully arranged networks of onsets on rhythmically different locations and metrically advanced divisions of the measure for all involved voices (ex. M.13-19), although also simultaneous onsets and abrupt changes such as sudden violent stops also occur (ex. M.56).

The section of *”Atmosphêres”* generally involves two different composition techniques: the building, sustaining and transformation of the vertical expression of tone clusters, as well as the organization of dense meso-textures based on *micro-polyphony*, which has been described by Steinitz as:

> ”...microscopic counterpoint, an internally animated yet dense texture in which large numbers of instruments play slightly different versions of the same line.” (Steinitz, 2003, pp. 103)

4.6.1.3 Orchestration techniques

The work is written for large orchestra and involves: 4 flutes (transverse flute as well as piccolo flute), 4 oboes, 4 clarinets, 3 bassoons, contra bassoon, 6 horns, 4 trumpets, tuba, pianoforte, 28 violins, 10 violas, 10 cellos and 8 double basses. The instruments of the orchestra are not divided into traditional voice groups but notated in individual *divisi* voices. This allows for a very large polyphonic potential that is exploited fully in large chromatic clusters - spanning up to about 6 octaves (ex. Letter ”B”). The timbral potential of the individual instruments are likewise exploited thoroughly by the use of both conventional and (at the time) unconventional playing techniques such as *con sordino*, *sul tasto*, many degrees of vibrato, *sul ponticello, punta d’arco, col legno, glissandi, harmonics* and *harmonic glissandi*, pitchless wind instrument sounds, stroking the piano strings with brushes and cloths according to carefully prescribed methods. The result is an immensely nuanced vertical expressiveness within the framework of a large symphony orchestra.

4.6.1.4 Basic techniques

”Atmosphêres” is devoid of the perceivable hierarchy of pitches associated with for example the tonal tradition and is without harmonic key and goal. There is almost no directional harmonic tension-release at play, although the systematic fading in and out of selected tone groupings does influence the experienced dissonance and harmonic entropy. Contrasts in the analyzed passage reside within register, register range, timbre, dynamics and micro-texture, which are facilitated most significantly

by polyrhythmic factors and choice of playing technique. The piece is in some parts very rich in rhythmic detail, but as is the case with the individual notes of the tone clusters it is difficult to distinguish the individual rhythmic forms from the complex polyrhythmic web they co-constitute because the meter is so consistently unarticulated (ex. M.44). The function of these complex rhythms, apart from contributing to meso and micro-textural structure, is an obscuring of the underlying meter (ex. M.18) and the creation of a collective rhythmic swarm (ex. M.24-28). The passage is governed by an absence of perceivable melodic and harmonic phrases, repetition, imitation, sequencing or other structural forms that can promote a sense of causal direction - in effect downplaying the significance of horizontality substantially.

4.6.1.5 Vertical models

First 29 measures - about 2:50 minutes

At the letter "B" a very large chromatic cluster appears spanning 6 octaves and a semitone including all the orchestra's instruments. From this giant total divisi tone cluster, and without any of the instruments changing their note, a series of instruments groupings are faded in and out starting in measure 18 with a group of wood winds and brass instruments that is playing the notes represented by the white keys on the keyboard. Subsequently, in measure 19 a pentatonic tone grouping is drawn out corresponding to the black keys of the keyboard played by flutes, clarinets and horns. Although this gives rise to a certain sense of progression of complex and highly entropic harmony there is strictly speaking no progression as all notes from start to end in these measures stay unchanged in the individual voices. Thus, what is presented is an effective variation in terms of dynamics and instrumentation that is capable of expressing movement in a strictly speaking harmonically static sequence, although the movement entails no causal direction and does not let itself predict. Furthermore, the technique resembles that of electronic music's subtractive synthesis in which filter banks extract tones and timbres from a noise source like white noise that contains all frequencies at equal amplitude. *"Atmosphêres"* here presents the orchestral equivalent to white noise with a total divisi chromatic cluster containing largely all the orchestra's notes from which Ligeti is then able to highlight, by orchestral filtering, the tone groups he wants[14].

Overall, the first 29 measures of *"Atmosphêres"* appear as being highly focused on the vertical rather than the horizontal axis. The section has no harmonic key, contains no motifs, repetition, phrases, melodies or pulse and exhibits no signs of causal direction. The excerpt develops through irreversible change in a continuous transformation of cluster based textures with no obvious note onsets. It is

[14] More than hinting at a connection between Ligeti's sound-mass techniques and his experiments with electronic music at Westdeutscher Rundfunk (WDR).

characterized by a high degree of musical entropy and the individual structures vanish in the collective whole of meticulously specified sub-elements.

4.6.1.6 Horizontal models

This is by no means to say that *"Atmosphêres"* does not exhibit horizontal musical structuring schemes, but they express themselves from "behind the scene" so to speak in the form of principles associated with serial techniques. The perhaps most interesting micro-polyphonic idea is that of an *intervalic concept* that gives rise to two tone rows that at first glance look like the sort of 12 tone series associated with basic serial techniques but that deviates from such techniques by having not the tone row itself, but the intervalic relations between notes as a foundation. This allows subtle irregularities to govern at the level of the audible pitch sequence while controlling these irregularities on the basis of highly rigid formalisms at the intervalic concept level. I will include here an example of such a compositional maneuver:

From letter "J" starts an intricate web of rhythmic events in the violins, violas and cellos over systematically arranged stepwise downwards and upwards tone movements in an overall chromatic but motion rich *divisi* cluster organized according to the following concepts:

Violins follow "Intervalic concept P" performing an overall downward direction

"Intervalic concept P"

|: [1⁄2 tone up] – [1 tone down] – [1⁄2 tone down] – [1 tone down] :|

… which produces the following tone row:

f#-g-f-e | d-eb-db-c | b-h-a-ab | g f-f#-e | d#-c#-d-c | h-a-bb-ab

… and viola and cello voices following an *inverted* version of the interval concept perform an overall upward direction

"Intervalic concept I"

|: [1⁄2 tone up] – [1 tone up] – [1⁄2 tone down] – [1 tone up] :|

… resulting in the following tone row:

c-db-eb-d - e-f-g-f# - g#-a-h-bb | c-d-c#-d# - e-gb(f#)-f-g

The tone rows arise as a consequence of ”Intervalic concept P” with the added rule of *inversion* at the level of the concept for each 12 notes (at the red line) giving the rows a palindromic structure. The opposing directions of the two instrument groups gradually narrow the cluster towards the middle register. And, as all of the *divisi* voices start at different places in the rows and exhibit different rhythmic behavior, the individual voices disappear in a collective myriad of perceptually inseparable internal movements - a micro-polyphonic canon with meticulous avoidance of any articulation of meter and a high degree of rhythmic, harmonic and melodic entropy. Yet, on a structural level, it is tightly and systematically organized based on a simple, formalized intervalic concept.

4.6.1.7 Summary - ”Atmosphêres”

The musical structure in the analyzed portion of *”Atmosphêres”* can be summarized by the following musical effects:

- Chromatic cluster harmony
- Very large dynamic range
- Unnoticeable onsets
- Very large polyphony
- Transformation of clusters through variations in range, register, timbre, dynamics and micro-texture
- Micro-polyphonic structures constituted by a large richness in detail in the individual voices and based on no more than a single palindrome intervallic concept and a rule of mirroring for every 12 notes, which is represented both horizontally and vertically
- Non-traditional playing techniques
- Slow gradual onsets and terminations of tone- and instrument groupings throughout
- Occasional highly abrupt changes

In other words, a number of musical aspects can be seen to deliberately have been avoided. These include: gesture, melody, pulse, articulation of meter and accentuated onsets.

From a phenomenological perspective, the music in the analyzed sections of *”Atmosphêres”* are characterized by a high degree of musical entropy on both a melodic, harmonic and rhythmic level. They present no audible causal direction, in so far as there is no tonal home, any melodic gestures or registrable repetitions, pulse, or other structural forms, which may be co-constituent of causal directionality. They give greater significance to the vertical dimension than the horizontal - e.g. by blurring metric perception. This is underlined by a very wide dynamic differentiation and nuance. In addition, the continuously flowing musical progression in both passages are experienced as *irreversibly changing* due to an absence of predictable musical goal direction of the vertically dominated expression.

Additionally, the passages clearly exhibit the characteristics mentioned earlier in this chapter as the three general characteristics for vertical domination by avoiding horizontally separable gestalt through the obscuring of onsets and terminations of notes as well as in the micro-polyphonic web facilitating a mass-merging of many individual events in to a mass texture. *"Atmosphêres"* also shows clear examples of avoiding perceivable horizontal regularities by an absence of repetition, rhythmic pulse and accentuation of meter as well as an absence of themes, motifs and rhythmic patterns. Lastly, it is evident that the analyzed passages of the piece show no causal direction and exhibit sections in which the period for event cohesion is exceeded.

4.6.2. STRUCTURE OF LIGETI'S ”LUX AETERNA”, G. LIGETI, (1966)

In *”Lux Aeterna”*, Ligeti uses similar compositional approaches as in *”Atmosphêres”*, however, some tendencies of vertical dominance are downplayed slightly in comparison.

Figur 9 - Screenshot of a lunar vessel flying over the surface of the moon in "2001: A Space Odyssey."

4.6.2.1 Filmic context

”2001: A Space Odyssey” features the first 3 minutes of *”Lux Aeterna”* playing from the beginning of the piece until measure 37 in a scene where a group of scientists are flying over the Moon's landscape in a small lunar vessel. The music seems to relate particularly to the frozen, grey and pristine landscape of the lunar surface as well as the contrasting view of the vessel and its human cargo's insignificant size compared with the surrounding space's infinite depth.

4.6.2.2 Overview - structure and meso-texture

”Lux Aeterna” (G. r. Ligeti, 1966) is a choir piece in 16 voices for soprano, alto, tenor and bass. The first section of the work (which is what is used in the film) is, however, written only for female voices and tenor. These are subdivided into 12 voice groups, 4 for each of the female voices and 4 tenor voices. The piece is marked with the words, *”FROM AFAR”* and opens *molto calmo.* The passage functions as a continuous flow without pause and without mentionable large changes. The overall form is characterized by a movement from a single note into a period of complex micro-polyphonic structure until it again fixates on a single note at the end (this time doubled by an octave). All voices go through the same tone row of 31 notes once in the duration of the section but start out staggered and run asynchronously and with different rhythms giving rise to a mutable meso-texture of

tightly woven counterpoint. Onsets are organized so that some voices exist rhythmically within a triplet division of the measure, some within a division of four, and yet others occupy each measure on a striation of 5 subdivisions, and onsets are marked *"all entries very gentle"*. Additionally, the ending of words that end on a consonant are notated to be left out in measure 36, which has the effect of avoiding any abrupt percussive sound in the otherwise uniform timbral development. In by far most of the section, the individual voices' performance of the tone row are not heard as independent melodies but, occasionally, melodic figures reveal themselves as, for example, in measure 12 in the 1st soprano with a hint of diatonicism in the otherwise anonymized micro-polyphonic counterpoint.

4.6.2.3 Orchestration techniques

The register range of the vocals are exploited to the fullest both in the bottom (e.g. 1. soprano's deep **Db4** M.9) and in the top (e.g. 1. Tenor's high **A4** M.24). Which gives a certain tension in the timbre at some places, for example towards the end of the piece where both sopranos and tenors are situated relatively high up in their respective registers. The harmony is closely spaced, has many voice crossings and is focused around the middle register. The passage as a whole evolves gradually from a focus on **F4** in the beginning through a maximum register range in measure 31 of an octave and a forth to **A5** with an underlying octave in measure 37. Nuance and evolution of the vertical expression happens at the level of timbre not least as a result of the different vowels in the text from the relatively low spectral centroid of the word "lux" over the slightly higher centered "aeterna" to the still higher "eis". Thus, a gradual opening of the timbre is taking place during the passage from beginning to end.

4.6.2.4 Basic Techniques

The closely spaced and continuous micro-polyphonic web of *"Lux Aeterna"* only to a limited extent uses structures that can be called motivic. The note material that constitutes the first three minutes of the piece only contains the notes included in the tone row and this row is followed slavishly by all voices. Rhythm is heavily syncopated and characterized by an alternation between longer legato notes and notes with relatively short durations. It never gives rise to accentuation, perceivable rhythmic repetition, rhythmic patterns or sense of musical pulse - partly due to the metrically highly varying onsets and the marking "all entries very gentle".

4.6.2.5 Vertical and horizontal models

Compared to the examples in *"Atmosphêres"* this excerpt of *"Lux Aeterna"* presents less chromaticism within the relatively narrow cluster structures, which arises as a consequence of close-spaced canon counterpoint - in turn dictated by the 31 note tone row, rhythm and metric distribution of entries. The tone row looks as follows (e.g. soprano mm.1-24):

f-f-f-e-f-g – f#-g-f-eb – ab-db-eb-f – gb-b-a-b – c-b-ab-f – g-eb-fb-eb – gb-f-b-g – a

There is no apparent underlying intervalic concept here. The lesser chromaticism is caused not least due to the tone row's makeup of both semitones as well as larger intervals, which, for the latter, entails the formation of "holes" in the chromaticism, even though the temporal distance between the staggered voice entries is short. This gives rise to occasional harmonically consonant vertical expressions, although by far the majority of the piece is quite substantially dissonant - an effect which is only further underlined by the harmonic density.

The piece also contains hints of melody. The most salient example is in measure 12 where the **b-a-b-c-b-ab-f** of the tone row stands out as a melodic gesture with a sense of diatonicism. Although there are occasional indications of consonant harmony and melodic gesture these occasions are without goal orientation and causal direction. Nevertheless, the moments of these structures function as hints of lower musical entropy in an irreversibly changing texture with an otherwise high degree of melodic and harmonic entropy.

The piece is, like *"Atmosphêres",* an example of how Ligeti utilizes very little musical base material - in this instance a single tone row - for a longer musical development organized through relatively simple principles.

4.6.2.6 Summary - "Lux Aeterna"

The musical structure in the analyzed portion of *"Lux Aeterna"* can be summarized by the following musical effects:

- Micro-polyphonic canon structure, in which all individual voices exhibit their own differing rhythmic and metrical characteristics
- Consistent, continuous flow throughout the piece
- Static dynamics
- The texture of the whole piece begins and ends in a one-note nucleus
- All voices are based on the same single tone row
- Details constitute the whole and disappear in it
- Evocations of melodic gesture, diatonicism and chord structures
- Chromatic as well as harmonic cluster structures
- Close harmony
- Many voice crossings
- Focus on the middle register
- Extreme vocal ranges give tension to the timbre
- Harmonic egality
- Gradual transformation of timbre and register throughout the larger form

Also in *"Lux Aeterna"* Ligeti seeks to avoid horizontally separable gestalts through obscuring of onset and termination and the mass-merging of many individual events. Perceivable horizontal regularities such as repetition, rhythmic pulse and accentuation of meter are avoided. The passage has no musical themes or rhythmic patterns as such, although some examples of motivic work do occasionally prevail. The bond between spectromorphologies is weak due to an absence of causal direction.

4.7. OTHER PIONEERS OF VDM

4.7.1. PENDERECKI

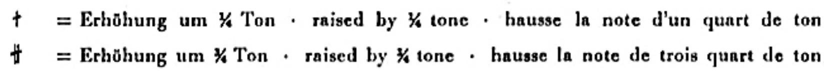
Similar compositional means have been used by the Polish avant-garde composer, Krzysztof Penderecki. Penderecki does not use micro-polyphony, but his cluster based works such as *”Threnody for the victims of Hiroshima”* (Penderecki, 1961) and *”Anaklasis”* (Penderecki, 1960) exhibit a clear focus on vertical expression in many passages and show the same three compositional characteristics outlined earlier and present in *”Atmosphêres”* and *”Lux Aeterna”* (i.e. voiding horizontally separable gestalts through obscuring onset and termination of spectromorphologies and mass-merging of many individual events; avoiding horizontal regularities through an absence of repetition, rhythmic pulse and accentuation of meter; weakening the bond between spectromorphologies through an absence of causal direction and by exceeding the period of event cohesion.)

Also Penderecki’s music has been used in cinema. It is featured in films such as *”The Exorcist”* (Friedkin, 1973), *”Inland Empire”* (Lynch, 2006), *”Fearless”* (Weir, 1993), *”Children of Men”* (Alfonso Cuarón, 2006) and *”Shutter Island”* (Scorsese, 2010), all of which have narrative aspects of mystery, something unknown, incomprehensible or horrific, paranormal activity or a presence of “something” that exceeds the boundaries of the senses. Most significantly perhaps, Stanley Kubrick’s hugely successful horror classic, *”The Shining”* (Kubrick 1980), featured no less than 6 pieces by Penderecki associating his music with scenes of horror and in situations where a connection to what could be termed ”the other side” is conveyed - effectively cementing his style in the *competence*[15] of audiences as accompaniment to not only space oriented science fiction (as in *”2001: A Space Odyssey”*) but also horror and mystery on film. As will be further elaborated in a chapter 6, the presence of VDM - as coupled to phenomena of outer space, infinity, mystery, paranormal activity and a presence of “something” that exceeds the boundaries of the senses - in the media competence of the player will help explain how the introduction of VDM in a game may shape the gaming experience.

[15] The notion of *competence* is described by Anahid Kassabian in regard to film as a skill that lets an audience *decode* a film’s intensions based on prior experience with the film medium. Competence will be discussed in chapter 6.

Penderecki is known also for his experiments in orchestral timbre and texture. This is evident not least in his abundant use of symbols and abbreviations to notate untraditional playing techniques. When examining a selection of his most used notational contributions it becomes evident that many of the sounds they are meant to produce are noisy, percussive or have to do with micro-textural expressivity and timbre. The most comprehensive list of Penderecki's use of such notational expansion can be found in the score for "*Fluorescences*" (Penderecki, 1962), which stands as a form of culmination of his experimental years (Schwinger, 1989, pp. 140). I want to highlight here, that these symbols expand the possibilities of vertical expressiveness in all main aspects of the vertical expression - that is, in regard to *tone system, harmony, timbre* and *texture*.

Symbols for the use of quarter-tones facilitate an expansion of harmonic expressivity by striating the octave into a 24-tone subdivided tone system instead of the traditional 12. This may be seen also as a timbral expansion.



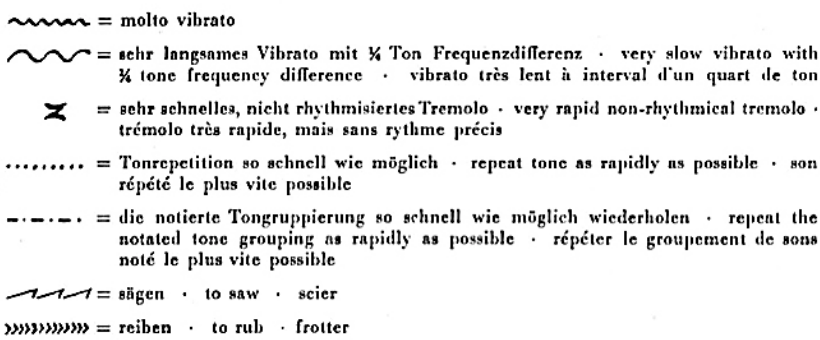
† = Erhöhung um ¼ Ton · raised by ¼ tone · hausse la note d'un quart de ton

‡ = Erhöhung um ¾ Ton · raised by ¾ tone · hausse la note de trois quart de ton

Figur 10 - Examples of quarter-tone notation in Penderecki's "Flourescences" (Penderecki, 1962)

Timbral expansions also include playing techniques for the strings such as *sul tasto* (on the finger board) abbreviated "***s. t.***", *sul ponticello* (near the bridge) abbreviated "***s. p.***" and *col legno* (with the wood of the bow) abbreviated "***c. l.***"

Other symbols facilitate a general expansion in terms of texture. These include *molto vibrato*, *very slow vibrato with 1/4 tone frequency difference*, *very rapid non-rhythmical tremolo*, *repeat tone as rapidly as possible*, *repeat the notated tone groupings as rapidly as possible* as well as *to saw* and *to rub*.



= molto vibrato

= sehr langsames Vibrato mit ¼ Ton Frequenzdifferenz · very slow vibrato with ¼ tone frequency difference · vibrato très lent à interval d'un quart de ton

= sehr schnelles, nicht rhythmisiertes Tremolo · very rapid non-rhythmical tremolo · trémolo très rapide, mais sans rythme précis

= Tonrepetition so schnell wie möglich · repeat tone as rapidly as possible · son répété le plus vite possible

= die notierte Tongruppierung so schnell wie möglich wiederholen · repeat the notated tone grouping as rapidly as possible · répéter le groupement de sons noté le plus vite possible

= sägen · to saw · scier

= reiben · to rub · frotter

Figure 11 - Examples of textural notation in Penderecki's "Flourescences" (Penderecki, 1962)

Other textural techniques include in the strings *several irregular bow strokes up and down in succession* and in wind instruments *frullato* - a whirring, very quick tremolo.

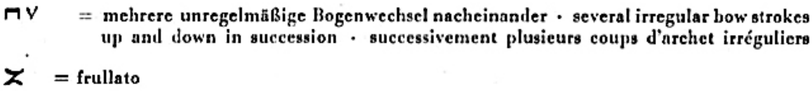


Figure 12 - Examples of additional textural notation in Penderecki's "Flourescences" (Penderecki, 1962)

Yet other symbols bear witness to an investigation into noise sounds like *strike the keys only* for wind instruments, as well as for strings *to be played on the tailpiece* and *to be played on the bridge* as well as different percussive techniques both on percussion instruments and on instruments that were not traditionally meant for percussive effects such as for strings *tab the soundboard with the heel of the finger-tip* and *tab the desk with the bow*.

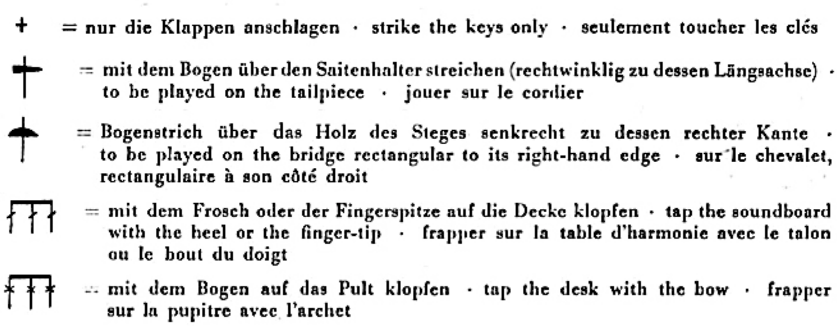


Figure 13 - Examples of noise sound notation in Penderecki's "Flourescences" (Penderecki, 1962)

Additionally, some symbols reveal a tendency in the composer to accept, to some extent, *unspecific notation* such as *playing the top note* of the instrument or playing the *bottom note* as well as the aforementioned symbols for frequent irregular bowing, rubbing, sawing as well as playing tremolos, repetitions and tine groupings as quickly as possible. Additionally, symbols for playing one, two or four strings *between bridge and tailpiece* can be seen to belong to this category of unspecified notational expansions.

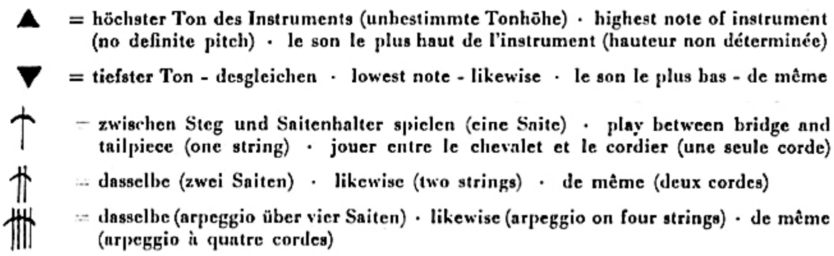


Figure 14 – Examples of unspecific notation in Penderecki's "Flourescences" (Penderecki, 1962)

4.7.2. XENAKIS

The Greek composer and architect Iannis Xenakis worked extensively with sound-mass music from the mid 1950s. Music that exhibits passages of clearly vertical dominance include: *"Metastaseis"* (Xenakis, 1953/1954) and *"Pithoprakta"* (Xenakis, 1955/1956).

"Metastaseis" for an orchestra of 61 players (1953-54) is Xenakis' first major work. It was first performed in 1955 at Donaueschingen Festival and features the concept of building musical masses out of a large number of events. Like much of *"Atmosphêres"* it is written in total divisi and presents systematic shaping of large glissandi clusters. (The piece was later to inspire the architecture of the famous Philips Pavilion in 1958). The intervallic structure, the duration of dynamics and timbres were conceived on the basis of geometrical ideals such as the golden ratio. (Bois, 1980).

A key compositional effect of the composer is the interplay of *continuous textures* (such as the unbroken glissandi clusters of *"Metastaseis"*) and *discontinuous textures* composed either from grains as in the tape based piece, *"Concret PH"* (Xenakis, 1958), where glowing coal produces a crackling sound-mass, or from fragmented instrumental gesticulations as in the orchestral work *"Pithoprakta"*. The latter work displays the composer's fascination with the relationship between the continuous and the discontinuous by developing gradually through the piece from point textures in the beginning towards continuous cluster structures as the piece progresses.

All of the above works clearly feature vertically dominated musical structures (in some passages more than others) by avoiding the formation of horizontally separable gestalts and perceivable regularities through a reluctancy towards repetition, pulse and accentuation of meter as well as weakening the bond between spectromorphologies by exhibiting no causal direction and by, on many occasions, exceeding the period of event cohesion. As is the case in the mentioned works by Ligeti and Penderecki, these works by Xenakis are characterized by frequently exhibiting what I have called *irreversibly changing static musical time* as well as *non-cyclic vertical time*.

Xenakis describes his methods as *stochastic*:

> ”In 1954, I introduced probability theory and calculus in musical composition in order to control sound-masses both in their invention and in their evolution. This inaugurated an entirely new path in music, more global than polyphony, serialism or, in general ”discrete” music. From hence came stochastic music.” (Xenakis, 1992, pp. 255)

Xenakis stochastic music is to a large extent based on mathematical equations and involves the use of probability as a key part of compositional process. These often computer aided works can, due to their probabilistic method, be seen to represent a middle ground between the determinism of the serialists and the indeterminism of contemporary composers such as John Cage.

Xenakis is known also for his interest in sound grains. The idea of sound quanta or sound atoms was investigated in the 1940s by physicist Dennis Gabor (Gabor, 1947), but can be traced back to the Greek atomistic philosophies of Leucippus and Democritus in the fifth century BC. Xenakis devised a compositional theory around granular sound included in his book, ”Formalized Music” (Xenakis, 1992). This work is later followed up by Curtis Roads, for example, in the already mentioned book, ”Microsound” (Roads, 2004). *Granular synthesis* is a technique in which sound *grains* of a short duration (i.e. At the micro time scale and approximately of the duration of what Daniel Stern refers to as *basic units of perception* ranging from 20-150 ms) produce masses of different timbres, micro-textures and densities.

4.8. SPECTRAL MUSIC - METHODS, CONCEPTS AND TECHNIQUES

Nowhere is the interest in the vertical axis as favored over the horizontal axis as in spectral music - especially in regards to *timbre* as defined in the previous chapter. Hence, a relatively large number of concepts will be covered in this section that have general relevance to sound at large as well as specific significance to spectral composition and to VDM at large. Most of the more technical concepts covered in this section of the chapter such as spectral envelope, instrumental spectra and attack transients are described in numerous sources on acoustics and psycho-acoustics. Joshua Fineberg presents such concepts in the context of spectral music. Where nothing else is stated, the terms and concepts explained in this section may be primarily attributed to two academic sources, both written by Fineberg: (Fineberg, 2000b) and (Fineberg, 2000a).

Spectral music - a term coined by French composer Hugues Dufourt - evolved in France in the early 1970s and was to become one of the most important schools of contemporary music in the 20th century. The primary composers of the style include Tristan Murail, Gérard Grisey, Hugues Dufourt and Jean-Claude Risset It can be seen as a continuation of the direction initiated by composers such as Debussy, Messian and Varèse where music is conceived as being made of sound rather than of musical notes. According to Tristan Murail, spectral music is to be understood as a compositional attitude rather than a set of techniques. It is an attitude of regarding music as sound evolving over time and seeks to compose music based on a refinement of our conception of sound; a new understanding of what sound is (Fineberg, 2000b). Spectral music is a way of externalizing intrinsic properties of sounds, about finding the essence of sound, enlarging that and composing with it. As the following pages will demonstrate, out of this attitude grew a broad pallet of compositional techniques.

Although based on an analytical focus aimed at spectral properties of natural sounds, spectral music is orchestral music, and its experimentation with electro-acoustic sound manipulation and mathematical models remains an influence and inspiration to orchestral composition rather than forming its own musical expression. In spectral music, properties within the frequency spectrum, such as vertical distribution and dynamic relations between overtones, are extracted from various sounds by Fast Fourier Transform analysis (FFT) and used as compositional material comparable to the way intervallic material is treated in traditional equal-tempered music. In other words, a very significant musical trait of spectral music is to emphasize *timbre* instead of tones, often rendering the distinction between harmony and timbre impossible. This leads Joshua Fineberg to suggest a form of vertical hybrid concept he calls *harmony/timbre* (Fineberg, 2000b, appendix 1, pp. 13). The musical content of spectral music is largely made of complex sound fields of harmony/timbre, which are combined in various ways and exposed to fluctuations of different kinds. The global musical form - its macro-form - is shaped by the evolution of these complex sound fields from one timbral state to another.

The vertical dimension is opened in spectral music. The striation of spectral space into 12 equal intervals is lifted and new orders are invented on the basis of overtone material, which is analyzed and manipulated. This music is very much based on new vertical paradigms and a significant portion of the spectral composers' work consists of inventing such paradigms through investigating the results of mathematical and electro-acoustic experiments on a spectral starting point.

The radical compositional approach posed by spectral music provides useful insight into the structure of the vertical. Not only does spectral music have clear tendencies towards vertical dominance, but the freedom it claims by occupying spectral space rather than pitch space, and the need for being much more precise in regard to vertical compositional parameters than 12-tone equal tempered music, calls for a special consciousness about vertical structuring paradigms, which suits the purpose of this chapter well. Therefore, a fair amount of effort will be made in this section to highlight these parameters. It should be noted that many of these parameters have traditionally belonged in the spectral composers' laboratories where they serve as models of inspiration for orchestral arrangements.

4.8.1. VERTICAL STRUCTURE IN SPECTRAL MUSIC

When we speak of the equal-tempered scale, we most often refer to the chromatic division of the octave into 12 equally spaced intervals. However, one might also conceive of equal-tempered scales with higher or lower division numbers - for example, 6 steps resulting in a whole-tone scale, 24 creating quarter-tones or 48, which makes eighth-tones. Greater subdivisions than 12 are defined as *micro-tonal*. While these micro-tonal scales increase the resolution of the vertical striation as the subdivision increases, their equal temperament still determines a fixed vertical grid, and although the precision of intonation is heightened, the stringency of the grid prohibits access to the physically infinite resolution of the vertical continuum. In spectral music, the use of quarter-tones and eighth-tones are based on necessity - from a wish to approximate, in an orchestral setting, spectral properties that are of the vertical non-striated continuum.

I will note here that the spectral composers' wish to utilize acoustic instruments and written scores despite all their efforts in the sound laboratories leads to an inaccuracy of pitch - this will become of some significance later in the thesis in regard to computer generated VDM as such limitations can be bypassed. A subtle distinction can be made in that in spectral music it is not artificial systems of organizing equal-tempered notes that form the basic underlying structure, but rather, these equal-tempered scales are the closest practical approximation of structures that belong to an infinitely nuanced vertical continuum. It is also interesting to note that to a wholly computer based generative musical output, whether through synthesis or sample playback, these approximations are unnecessary and the unstriated, smooth continuity of spectral space can be preserved, while at the same time allowing orchestral instrumentation.

4.8.2. ORCHESTRAL SYNTHESIS

One of the basic methods in spectral composition is *orchestral synthesis*, where the principle of additive synthesis is transferred to an acoustic ensemble (Fineberg, 2000a, pp. 3). In additive synthesis, timbres are constructed out of simple sine waves through the addition of overtones to a fundamental pitch. The characteristics of the timbre are controlled by mainly two parameters; the number of overtones and the amplitude relations between partials. The vertical distribution of partials in the *harmonic* spectra is naturally fixed. The harmonic partials produced by simple multiplications to a fundamental frequency are unique to that fundamental and the auditory system is so attuned to this harmonic series that we will perceive the fundamental frequency psycho-acoustically even in situations where it is not physically present and its overtones are the only actual physical stimuli. The simple multiplication ideal means that even when the pitch used as the fundamental frequency falls within the chromatic equal-tempered scale, many of its overtones do not. This, in the context of orchestral synthesis, produces harmony/timbres that go far beyond their natural sound modeling prototypes in terms of sonorous complexity as the acoustic instruments add their own overtones to the mix. In addition, the simple equation of harmonic partial distribution mentioned above must also be seen as an ideal. In reality, when pitches are being produced by real instruments, a number of acoustic and physical factors contribute to the spectral profile becoming significantly more complex. For example the presence of noise from the bowing of strings, wind instrument breath sounds, instrument formants, spectral envelopes and attack transients as well as the overtones that arise from the principal waveform of different instruments (bowed strings have triangle-like waves, clarinets have square wave like shapes and so on). The notion of *instrumental spectra* is a significant timbre-shaping reality, also in spectral music.

4.8.3. PARAMETERS OF SPECTRAL MUSIC

Formants are filtering effects caused by physical and acoustical properties of a sound producing, vibrating body. The formant profile of any sound source is unique to that source or group of sources (female voices could constitute one such group), and this is why formants are also a primary catalyst for human identification of sound source into specific voices, instruments, animals and so on. In the example of language, different vowels are produced as a result of formants created by the variable shape of the mouth.

In the same way that a dynamic envelope of a sound object refers to the evolution in amplitude over time of that sound object as a whole, *spectral envelope* refers to the relative amplitudes of partials as these relations evolve over time. As is the case with formants, each sound source or group of sound sources will have certain common characteristics in regard to their spectral envelope, which set them apart from one another in our perception. Micro-variations in the spectral envelope, or *spectral flux,* is a key component to natural sound and is why spectral composers have traditionally mimicked such variations in their work within the orchestral setting.

The term *attack transient* applies to the often noise-like and relatively loud initial portion of a sound. (That is very important in regard to timbre recognition as discussed in chapter 2.) I will argue that attack transients, if they are obscured, therefore represent a potential powerful means for listener alienation towards a sound. This goes along the same aesthetic lines as other principles of VDM as we saw with Ligeti and Penderecki. In spectral music, the attack transients are rarely mimicked precisely, but are implemented based more on intuitive artistic choices.

Our capacity to perceptually fuse partials, and thereby segregate a sound source from its surroundings, is, among other things, dependent on these partials being harmonic. The strict order of partial harmonicity described in chapter 3 is a natural artifact of most pitched sounds whether instrumental or non-instrumental.

Non-harmonic partials, however, also occur in the orchestral setting in the form of bells and cymbals and the distribution possibilities of these partials are infinite.

These sounds do not fuse to a single pitch and timbre in the auditory system but provide a more noisy, complex sonic expression. Both harmonic and non-harmonic spectra can be designed by spectral composers, creating new sounds either tangent to or alien from naturally occurring sounds. One method for creating models that produce these new sounds is additive synthesis. By applying simple mathematic equations to the partial distribution an effect of *spectral compression* and *stretching* can achieved whereby harmonic overtones can be made to either grow further apart or contract in relation to one another. Another type of artificial distortion of the overtone series is by *spectral pitch shifting*. This is done by adding a fixed value in Hz to the frequencies of all partials .

Different effects can be obtained by using *amplitude, frequency* or *ring modulation* as manipulation of a carrier oscillation. Amplitude modulation (AM) can produce tremolo effects (amplitude vibrato) but when implemented more subtly can mimic the ever-changing micro variations in amplitude of natural sound, enriching a sound's micro-texture. Frequency modulation (FM) produces variation in the pitch of a carrier oscillation caused by a modulator oscillation. As with AM, the rate and depth of the modulation can be controlled via the frequency and amplitude of the modular wave as it effects the carrier. FM above 20Hz produces symmetrical overtones and undertones centered around the carrier frequency. Thus, providing many oscillations with only two oscillators, makes frequency modulation a very efficient technique for creating a wide range of artificial timbres from relatively limited resources. This is a technique exploited by early sound chip manufacturers in the game console business (see chapter 7). Ring modulation produces the sum and difference of two signals, creating its own distinct electro-acoustic effects, which can then be simulated orchestrally.

4.8.4. RESERVOIRS

The above mentioned components in the work of spectral composers can be used to create *reservoirs* of model material - underlying structures that can be referenced when composing much like harmonic key and different modes are used as models in tonal composition or the primary tone row in serial techniques. As with modes and, sometimes, tone rows, such reservoirs can facilitate coherency in the musical expression. It is important to note that these reservoirs can take shape in many ways. They can be mathematical relations ready to be implemented on sound for later orchestral synthesis or specific calculated spectral properties.

4.8.5. PRINCIPLES OF INSTANTANEOUS GENERATION

Gérard Grisey uses the concept *principle of instantaneous generation* (le principe de la génération instantanée) (Moscovich, 1997, pp. 25), whereby one sound *component* - (i.e. one characteristic of sound such as a fluctuation pattern, harmonic distribution, partial intensity ratio and so on, brought about by manipulation and/or analysis) - is selected, used compositionally, and from which new sound-entities can be chosen. The musical states are formed as a chain-reaction so that each new state is based on properties of its predecessor.

This sounds a lot like causal direction, although of a very un-orthodox kind, where it is not melodic relations and harmonic progressions that are evolving with causal direction, but rather textural and harmony-timbral effects of imposing mathematical calculations and electro-acoustic manipulation to a spectral starting point. An important point to notice here in regard to my endeavor to generate music automatically, is that in Grisey's principle of instantaneous generation, these sound components, which determine the characteristics of the next sound-entity in the chain, are subject to an aesthetic choice made by, and notated manually by, the human composer.

This *becoming* of new sounds *(devenir des sons)* is the core in Grisey's method and aesthetic in which:

> ”Musical form becomes the projection of natural microphonic space into an artificial screen which serves to deform, to focalize, to multiply, to select, and so on.” (Moscovich, 1997, pp. 25)

The connection to *vertical dominance* as defined in chapter 2 is very apparent. Grisey writes:

> ”By the attention brought, constantly, no longer on the material itself, but on emptiness, on the distance which separates one perceived instance from the next one (degree of change or evolution), I think I have come a little closer to the fundamental time, no longer chronometrical time but the psychological one and its relative value.” (Moscovich, 1997, after M. De la Cruz Padron Lopez De-Lilly, idem., pp. 75, et 79-81)

4.8.6. SPECTRAL MUSIC - A CRITICAL NOTE

A problem in orchestral spectral music, its perhaps most significant challenge, I argue, is the inaccuracy with which analysis results are translated to musical notation for live musicians in a symphonic setting. The temporal quantizing needed and the harmonic-timbral approximations dictated by whichever micro-tonal system is applied detaches the sounding music from the natural ideal on which it was conceived. The spectral phenomena extracted from acoustic sound are quite significantly obscured in this process due to obvious audible effects of the fine vertical inaccuracies and the addition of instrumental spectra. Strictly speaking, it would seem that spectral music, through its orchestral form, has a certain integrity problem here. However, just as a certain flexibility is allowed in the translation of analysis results into musical notation, so, too, does flexibility reside between this notation and the performing musicians offering the possibility for the ensemble to intonate according to ability and in the context of the sonorous reality of the music as it is played in a concert venue. I will go deeper into such questions when examining what I have called ”music-shaping stages” in chapter 8.

These considerations are interesting in regard to computer-generated, computer-performed music. Although far from all software samplers on the market offer any micro-tonal control, it is perfectly possible to build a sampler with perfect pitch precision, thereby making possible the exact intonation required to more precisely mimicing the natural analyzed sounds, while at the same time allowing orchestral instrumentation. Certain technical and aesthetic challenges arise from this task - such as the very close proximity of many micro-pitches tending to draw on the same samples, if one does not rethink the idea of the traditional sampler and sample library as a means for manifesting generative music. Such close spectral proximity of identical samples played at minutely differing speeds can result in a lack of divergence in the sound as well as possible phase problems.

The challenges of computer-generated music is discussed in chapter 8 and further elaborated in Appendix A and B, where practical experiments with realizing VDM via a generative algorithm and a costume built sound library are presented.

4.9. OTHER STYLES WITH VDM ASPECTS

The music-structural characteristics of VDM are not limited to the sound-mass music of Ligeti, the Polish sonorism associated with among others Penderecki or the stochastic and granular methods of Xenakis. Examples of the general characteristics of vertical dominance are found in a range of musical styles, which I will not go deeper into here. La Monte Young, Morton Feldman and Arvo Pärt have all flirted with varies degrees of vertical dominance as have a vast number of other contemporary composers. Indian tambura music, drone music, noise music and some ambient music may also contain these traits to various extents. Even some examples of the black metal genre can be seen to exhibit vertical dominance because of its often extremely fast playing and noisy production, which fuses events together to a mass.

4.10. CONCLUSION

The historical development of verticality in composition represents a gradual liberation of musical structure from being governed by strict rules within pitch space to freely occupying spectral space. VDM, as an umbrella term, can be associated with music across styles and historical periods on the basis of common underlying structural and perceptual properties. These general underlying principles of VDM and a range of vertical parameters, to add to the list of parameters presented in the previous chapter (and listed in Appendix F), serve, in the larger context of the thesis, to provide a basis for understanding the structure of VDM, and in turn allow the formulation of an initial set of compositional principles towards the end of the thesis.

The determining factors lie in the presence of certain musical traits and, not least, in the absence of others. In regard to the latter, three general characteristics that facilitate vertical dominance were identified in this chapter:

- **Avoidance of horizontally separable gestalts**
    - By obscuring of onset and termination
    - By mass-merging of many individual events
- **Avoidance of perceivable horizontal regularities**
    - Due to an absence of repetition, rhythmic pulse and accentuation of meter
    - Due to an absence of patterns like themes, motifs and rhythmic patterns
- **Weakening of the bond between spectromorphologies**
    - By an absence of causal direction
    - By exceeding the period of event cohesion

In addition, a set of techniques for the expansion of vertical expressivity has been listed and includes:

- New orchestral playing techniques
- Emphasis on mass-effect rather than individual pitches and rhythms
- New musical texture conceptions
- Experiments in timbre
- New orchestration techniques
- New technologies
- New horizontal development schemes
- Focus on the perceptually *unspecific*

I have defined *musical entropy* on the levels of harmony, melody, timbre and rhythm. This is a useful term in describing structural characteristics of VDM, not least as the degree of musical entropy may be tied to the degree of vertical and horizontal dominance - an aspect of VDM perception that will be elaborated in the next chapter. I introduced the notion of the *rhizome* as a metaphor for VDM structure due to its *non-hierarchical, a-centered* and *non-linear* characteristics that resemble those of auditory verticality as opposed to arborescent, tree-like structures, which represent hierarchy, causality and an underlying generative model.

Works by Ligeti, Penderecki, Xenakis, Grisey, Murail and a range of other composers and styles exhibit vertical dominance by complying to the three general characteristics of VDM as well as utilizing some of the expansions of vertical expressivity mentioned in this chapter. A non-comprehensive list of VDM methods includes Ligeti's micro-polyphonic technique, Penderecki's experiments in playing techniques and Xenakis' interest in probability as the basis for music composition as well as the relationship between the continuous and discontinuous textures through, for example, granular synthesis. Furthermore, the use of Penderecki and Ligeti on film cements their VDM styles with space oriented science fiction, horror and mystery on film giving hints to how the music may influence the gaming experience if implemented in a game.

Spectral music offers a range of ideas and music compositional approaches - some of which can inspire structuring schemes for generative music creation. Being the perhaps most radical advocate for vertical dominance, spectral music utilizes vertical parameters such as: *micro-tonality, orchestral synthesis, instrumental spectra, formants, spectral envelope, spectral flux* and *attack transient* in its composition methods of *orchestral synthesis.* Other techniques associated with spectral music include: Spectral stretching and compression; spectral pitch-shifting; amplitude-, frequency- and ring-modulation; amplitude- and frequency-modulation as a parameter for micro-textural enrichment of samples. While amplitude and frequency modulation have both been mentioned in chapter 3, I will add *ring-modulation, spectral stretching and compression* and *spectral pitch-shifting* to the list of vertical parameters initiated in the previous chapter and presented in Appendix F.

Reservoirs have traditionally served as a pool with different approaches to vertical striation consisting of relational information, which can act as a blueprint for structure creation. Such reservoirs have included various micro-tonal configurations based on partial distribution and dynamic relations that are extracted via FFT analysis of chosen acoustic sounds and electro-acoustic manipulations.

I have introduced the idea that the concept of reservoirs holding timbral properties for compositional exploitation may also be utilized to create textural templates that can be superimposed or co-create vertical expressions. Additionally, in regard to the broader scope of VDM, the idea of vertical reservoirs would not have to necessarily confine itself to spectral striations but could also include other vertical striation paradigms such as equal-tempered chromatic tuning, diatonic, pentatonic, church modes and other scales. Orchestral synthesis' inaccuracy of intonation, I have argued, is at the same time a challenge and a defining factor of the style. In the context of entirely computer-based VDM composition and playback, orchestral synthesis holds the potential to become pitch-perfect without leaving behind the orchestral instrumentation.

4.10.1. SIGNIFICANCE OF SPECIFICITY

A common denominator in the VDM styles presented in this chapter is a tendency towards perceptual un-specificity. The significance of specificity in regard to the detail of the sound-mass seen as a meso-texture is relatively small from a perceptual point of view compared to the expressive impact of larger form components and to more general structuring schemes. In other words, the sound-mass as a whole is more significant than its constituting parts. This is a general theme across sound-mass composers - examples can be found in Penderecki's use of symbols for unspecific playing techniques and in Ligeti's complex polyrhythm in which the idea that no onsets may happen at the same time seems perceptually more significant than the exact timing of individual entries - although Ligeti, on a structural level, is highly accurate in his notation. Xenakis' use of probability as compositional method goes along the same lines. The spectral composers seem to be dependent on an accuracy in the performance of their notation in order to accomplish their delicately nuanced harmony/timbres, although the inherent inaccuracy of the orchestral synthesis concept can itself be seen as an expression of un-specificity.

Also the sequential order of notes in a voice is less important in a dense sound-mass structure than compositional choices based on vertical parameters. A key tendency of these composers in their sound-mass works is on one hand to have an impressive detail in the written score, (not least out of necessity because of the involvement of musicians who have to have something concrete to abide to) while, on the other hand, using this detail to accomplish textures in which those details are in and of themselves largely insignificant to the overall musical expression. The immense detail with which micro-polyphonic structures are conveyed to the musicians is a necessity when working with an orchestra. In other words, there is a need for specificity because of the presence of an orchestra. If one takes away the orchestra, this need disappears with it.

Across acoustic VDM styles, a tendency to challenge and expand the existing frame of the traditional orchestral paradigm (including its written score, performance, spatiality, and musical invention) in order to obtain a desired vertically dominated musical expression is clear. Penderecki's and Ligeti's notational expansions is one example, the spectral and electro-acoustic traditions' transgression of pitch space another and the use of extreme playing techniques as well as both ultra-complex and unspecific cluster notation are yet other examples. This, together with a general aim to obscure and hide gestural and melodic structures (which in Penderecki and Ligeti's works calls for extreme caution and precision in the avoidance of synchronicity) in order to produce sound-masses of vertical expression, did not, however, break out of the orchestral paradigm, but introduced new directions within it.

Generative music in games, where music is being composed, ”interpreted” and played by the same system, allows emancipation from many of the requirements connected to the traditional orchestral paradigm. These include the need for a written score, a need for individual voices and a closed musical macro-form; the latter being largely irrelevant due to the non-linearity of the game media - though not fully irrelevant as will become evident later. In such a compositional scenario, which can be synthetic or based on audio samples, breaking out of the restraints is made possible - importantly, without necessarily leaving behind the orchestral sound.

(There are also very significant limitations connected to producing orchestral music entirely in the computer, these will be taken up in chapter 7 and 8.)

I do not mean to say that the acoustic VDM mentioned in this chapter is in any way incomplete or imperfect. Its expressive resolution within the orchestral frame it is bound to is infinite due its acoustic instrumentation and human interpretation, while computer-generated VDM from this perspective is limited by its status as a form of complete design, devoid of any randomness other than the one implemented on purpose as part of that very design. This topic, as mentioned, gives rise to an examination of music-shaping stages in chapter 8.

The findings of this chapter together with findings of the previous two chapters will serve as part of a theoretical basis for the development of a broader theoretical model of the relationship between musical structure and musical listening states. This is the subject of the next chapter, which is concerned with the perception of VDM.

CHAPTER 5. PERCEPTION OF VDM

Having defined what constitutes the perceived vertical, I was able to list a series of structural circumstances that constitute VDM and that influence the vertically dominated musical expression. It is now time to look at a few prominent perceptual issues regarding the experience of VDM and music more generally. In the present chapter the case will be made that general VDM structuring schemes, discussed in the previous chapter, can be seen to facilitate a special listening state, *now-listening*, which differs from other musical listening states at the level of musical anticipation. Furthermore, when the experience of music is seen as tied to embodied perceptual capabilities facilitated by the human mirror neuron system, a set of interesting relationships can be identified between the structure and perception of VDM and music in general.

5.1. INTRODUCTION

The focus in the present chapter on the perception of VDM stands as an important step to the thesis' aim to present a phenomenological and music-psychological framework of understanding, which, together with an examination of VDM's narrative potential in the next chapter, can form the basis for determining the impact of VDM on the perception of a game or other interactive experience. The general perceptual attributes of VDM presented here serves to inform the clarification of this narrative potential. It does so by highlighting the relationship between VDM structure and perception from the perspective of musical anticipation as well as taking the consequences of seeing *musical listening* as what I will refer to as *empathetic co-agency* - a musical enacting of motor-functional, linguistic and emotional with a neural basis in the human mirror neuron system.

The chapter is thus divided into two main sections. The first one deals with presenting a model for music perception that is based on musical anticipation. I call this model *Three anticipatory listening positions.* The second section argues for the perspective on musical listening as empathetic co-agency and investigates the implications of such a perspective in regard to the perception of VDM and music more generally. The two sections together allow the formulation of a set of perceptual attributes of VDM that can be explicated in a list of adjectives useful in the further treatment of VDM's narrative potentials of chapter 6.

As an opening to the anticipatory model for VDM perception in the section called "Three anticipatory listening positions" a short view at Husserl's conception of a three part present moment is helpful and invites the terms *retention* and *protention* into the discussion. Jean Piaget's widely assumed notion of cognitive schemas as a model for human learning is discussed and concepts such as cognitive *equilibrium*, *perturbation* as well as the subject's adaptation to its environment through *assimilation* and *accommodation* is covered. *Historical musical form* is seen by Ligeti as the total musical past of a listener - encompassing both veridical, schematic as well as long term memorized musical experiences into a single category. This category serves as a naming factor for an omnipresent listening position in the anticipation model, *historically-based anticipation*. After an initial presentation of its theoretical background the chapter thus goes on to elaborate the three anticipatory listening positions comprising *equilibrium-based anticipation* (or future-listening), *historically based anticipation* (or past-listening) and *perturbation-based anticipation* (or now-listening). These positions are tied to concrete musical structuring schemes of the previous chapter. The model acts as a theoretical cornerstone in regard to the perception of VDM - in particular the connection between the structural characteristics of VDM and perturbation-based anticipation or now-listening are of special significance to the thesis. The anticipatory model does not confine itself to VDM, however, but allows for the positioning of this special musical style in the context of musical anticipation more generally shedding light also on the perception of HDM.

In the section of the chapter called ”Music as empathetic co-agency” I propose a definition of music as *that* which is listened to in a certain state of listening, *musical listening*, which involves *empathetic co-agency* - a term that, as mentioned, combines linguistic, motor-functional, musical and emotional aspects of mimicry allowed by the human mirror neuron system. By presenting a brief overview of neuroscientific evidence in the field a plausible link between perception and action facilitated by the human mirror neuron system is highlighted. Some studies of this overview also suggest a plausible link between the perception and action involved with such modalities as motor-action, language, emotions and music. Based on this, I argue that the experience of music as tied to these other modalities through the mirror neuron system as a shared neuro-functional substrate may be seen as a foundation for the understanding of *musical listening* as *empathetic co-agency.* Empathetic co-agency is in turn dependent on various degrees of hierarchical structures, patterns and musical horizontality and may be tied to various degrees of experienced *sense of control*. I argue that on a music-cognitive level, as well as at the level of embodied experience, a movement towards now-listening in the model (which VDM per definition performs) is associated with a sense of control loss, whereas moving towards future-listening is associated with a sense of gaining control. On this background I go on to argue that the *vertical expression* seen as an auditory *singularity* is not experienced through *musical listening* - and is therefore from a music-perceptual perspective not associated with perturbation, but rather equilibrium - in effect tying the extremes of the model together into a circular model.

In the subsection, ”The undisclosed sender of VDM’s empathetically co-enacted intentions”, I philosophize that the question of an undisclosed, unidentified, non-concrete and perhaps illusory *sender* (in a linear sender-medium-receiver communicational paradigm), who takes the role as an origin for *intentions* and is thereby occupied with a certain *state* with which the receiver can co-enact empathetically, may be present intrinsically in the musical listening state. This leads to the crucial question: what is the state of an undisclosed sender whose motor-functional, linguistic and emotional properties when mediated through sound produce VDM? I propose that the perception of VDM can be partly understood as the perceiver's empathetically co-enacted "answer" to this question - in the form of a mental representation based on mimicking capabilities of the mirror neuron system that facilitates a realization of the sender's state through music listening. I further argue that this role of sender of *intentions* may be attributed by the perceiver to narrative elements of a multi-modal percept - thus audio-visually linking non-diegetic music and diegetic narrative elements like characters, avatars, the presence of "something" that exceeds the boundaries of the senses or the narrative setting such as for example outer space.

In the subsection, "Perceptual attributes of VDM - the state of the sender", a list of adjectives is presented (continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless and unpredictable) that may be attributed to the perception of VDM based on perceptual attributes of perturbation-based anticipation as well as the perspective on musical listening as empathetic co-agency.

In the subsection, "Movement at the musical event horizon", I suggest that VDM receives its special perceptual qualities by situating the listening position very close to the border of the vertical expression (the event horizon) and by constantly crossing the border and coming back out at the same time maintains a state of musical listening while constantly violating expectations.

Finally, I will discuss emotional attributes of VDM in regard to positive and negative valence and argue that it is primarily the vertical expression and the horizontal development of it that sets the emotional agenda. However, as a highly significant secondary attributor, the perceptual attributes of vertical dominance discussed throughout the chapter (tied to the adjectives mentioned earlier) change the quality of such emotional attributes of the vertical expression by adding their own unique qualities.

5.2. THREE ANTICIPATORY LISTENING POSITIONS

5.2.1. RETENTION AND PROTENTION

Husserl's model of time-consciousness (Husserl, 1966 [1887]), which has influenced philosophers such as Heidegger, Sartre and Merleau-Ponty, maintains that the perceived present moment has duration and can be divided into three internal parts: a past-of-the-present-moment that he calls *retention,* a future-of-the-present-moment he calls *protention,* and a present-of-the-present-moment (Stern, 2004). Retention can be understood as a fantasized memory of the immediate past held in the present moment, while protention is an anticipation of the immediate future likewise held in the present (Gell, 1992, pp. 223). Both retention and protention are thus parts of the present moment, but their content, so to speak, is tied to past and future respectively.

I do not wish to submit to the theoretical context in which the terms are conceived as much as to merely utilize them for their clarity in terms of describing the immediate past and future as situated within the experienced presence, as well as distancing the discourse from seeing the present moment as a knife's edge between past and future with no duration or content of its own. With a basis in such a three-part conception of the present, I thus suggest in this section *three anticipatory listening positions* that are tied to these three elements of Husserl's model of time consciousness through their degree of musical predictability.

The model combines, among other aspects, the already covered concepts of *musical entropy* and the phenomenon of *irreversible change* with Piaget's widely used notion of *cognitive schema* as well as Ligeti's conception of *musical history.* Before I go into an overview over this anticipatory model, I will therefore first take a look at the two latter notions.

5.2.2. COGNITIVE SCHEMAS

The Swiss cognitive scientist, Jean Piaget (1896-1980), introduced the idea of cognitive schema in his endeavor to understand the psychological development of children. Since then the concept has gained ground in many areas other than children's development and is broadly recognized within cognitive psychology to be of a somewhat universal character and applicable in understanding thinking also in adults (e.g. Glass et al., 1990). Considerable critique has been raised to the larger theoretical body in which the notion of cognitive schema are a part since its conception (on the subject of children's development), why I also here wish to distance myself appropriately from such theoretical context and subscribe merely to the widely acclaimed usefulness of the term at hand.

The cognitive schema can be explained as a cognitive structure that the individual builds over time from knowledge about the world based on past experience. This happens continuously as a fundamental process through which the individual understands. We build cognitive schemas for everything we experience in the world around us. Constituents of a schema for a school would thus include teachers, students, books, and so on, not airplanes and tourists. These would normally belong to a schema of an airport. A schema of Viennese classical music is constituted by for instance functional harmony, strings and sonata form, not a sweating audience, guitar riffs and double bass drum that may be connected to a rock concert. But schemas operate on many different levels and also include our unconscious understanding of culture, other people, self image and so on. Regarding the individual's understanding of the world as cognitive structures is a fundamental pillar in cognitive psychology. The cognitive schemas are, however, revised all the time in a constant process of *adaptation* between the world and the individual's understanding of it. This adaptation can occur through *assimilation* or *accommodation* (e.g. Hauge & Thomsen, 1999, pp. 133). By assimilation new experiences are incorporated in an already existing schema. If the existing schema and the new experience are incongruent a cognitive imbalance arises, which is often referred to as *perturbation*[16]. By assimilation, then, the experience is adapted to supporting the existing schema, ensuring a maintenance of cognitive balance. This balance is referred to as cognitive *equilibrium.* The individual's basic strive towards equilibrium and thus away from perturbation is the fundamental driving force which leads to either assimilation or accommodation. *Accommodation* is to be understood as a somewhat opposite process, where in this strive for equilibrium it is the individual's cognitive schema and not the experience that is adapted as a result of incongruence between experience and schema. In the case of many experiences, which do not fit to existing schemas, the schema is rejected in favor of a new schema, a new stage, which is capable of accommodating the new experience. Stage by stage, in a reciprocal interaction between assimilation and accommodation and driven by a basic strive towards equilibrium, the individual develops an increasingly accurate understanding of the world. An extensive study of among other things schematic expectations in regard to the experience of music may be found in David Huron's book "Sweet Anticipation" (Huron, 2006).

5.2.3. MUSICAL PAST AS HISTORIC FORM

Geörgy Ligeti puts forward an interesting perspective on our perception of currently sounding musical form in the light of *historic form*. Musical form is identified by Ligeti as a spatial abstraction of music, where musical events in their extension in time appear as shapes in an imaginary space. Ligeti writes:

> "…"music" should thus be the purely temporal event, "musical form" on the other hand, the abstraction of the same temporal duration, in that the relations within the duration are not presented temporally but virtually spatially; musical form occurs only when in retrospect one overviews the duration of the music as "space". Retrospective overview means history. […] Both the immediately played music's earlier moments and all prior experienced music is present in the musical form." (Translated from German, G. Ligeti, 1966, pp. 47)

Ligeti's conception of historic musical form, argued here, is relevant because it denotes the implications of prior musical experience on what is currently heard. Thus, this historical aspect comprises both the immediately heard in the listener's

[16] Perturbation is not an expression coined by Piaget but is used broadly in psychology as an expression of uneasiness or disturbance.

short time memory, which is part of immediate musical context or musical piece, as well as the listener's musical history with all that this entails of personal preferences, available music, cultural traditions etc. Here Ligeti's conception of *history* builds on a dynamic understanding of time in that formal components in the past are connected to components in the present as subdivisions in a larger form that includes the total prior musical experience of the individual, and as such is an undivided, continuous process through which the individual continually develops its unique premise for music perception. The article from which the quotations are taken, is characterized by a certain categorical approach, where the separation of music and musical form as something that only occurs "when retrospectively overviewing the music's process as "space", may seem in some sense to decouple the musical form from a perceived immediate sounding musical reality. The anticipation of form is based on prior experience and hence on history. This is not a new idea, and wasn't either at the time of writing in 1966. It seems to be a universally valid and widely accepted principle in academic musical discourse that the musical past - in particular within the frame of the same piece - has an impact on the experience of the musical "now".

The reason why it is relevant in the context of the anticipatory model posed in this chapter is that a *historically founded anticipation,* albeit always active in a certain sense, steps into the background as the most significant anticipatory listening position, when: The listener cannot predict the further musical development and thus is unable to use information from the past to form an informed guess, or, when the listener knows exactly what's coming and therefore relies on projections of the past into the future. I use the word *history* in this meaning and place *historically based anticipatory listening* as a field between two extreme positions, *equilibrium-based* and *perturbation-based anticipation*, maintaining that musical listening in the present moment can never be completely devoid of neither musical retention nor schematic, veridical[17] or historical expectation.

5.2.4. OVERVIEW OF THREE ANTICIPATORY LISTENING POSITIONS

Focusing on musical anticipation, the model of three anticipatory listening positions presents three perceptual modes relevant to different musical listening experiences. I have, as mentioned, chosen to call these modes *three anticipatory listening positions.* The word *anticipation* contains, however, at least two different meanings: *prediction* and *expectation.* Where expectation can be regarded as passive in its

[17] Bharucha in (Bharucha, 1994) distinguishes between schematic and veridical expectation in music. Schematic expectation is seen as culturally generic, automatic and related to a general schema of how music in general behaves, while veridical expectation is based on the specific musical piece being heard. Ligeti's notion of historic musical form in this respect relates to both general schema and specific veridical expectation.

relation to some future occurrence, prediction is active. Additionally, ”expectation” seems associated with a degree of uncertainty, while ”prediction” to a larger extent carries a sense of certainty about the future. It is precisely the ability to harbor both of these meanings that makes the notion of *anticipation* useful in the context of the model. This will become more apparent in the following pages.

The model represents a continuously variable scale, a field in which every musical listening situation could be placed according to degree of musical predictability and associated with certain cognitive implications. The scale ranges from what I have termed *equilibrium-based anticipation* over *historically based anticipation* to *perturbation-based anticipation.* Which of the three anticipatory listening positions exist in a given listening situation is determined through universal cognitive preconditions of *assimilation* and *accommodation*, which are co-constituents of Piaget's notion of the cognitive schema. Moreover, it is determined by the listener's cultural position and knowledge of the music as well as a defined set of specific musical qualities, including the degree of *musical entropy*, presence or absence of *causal direction*, *horizontal* or *vertical dominance*, the amount of *repetition* and whether musical events occur *cyclically* or *irreversibly changing*. These concepts are discussed in detail in a moment.

The framework of the model is governed by three general aspects. The first aspect accounts for *specific musical qualities*. These include oppositions such as high/low musical entropy, variation/repetition, horizontal/vertical domination. The second aspect is concerned with the *universal cognitive pre-conditions:* Equilibrium/perturbation and assimilation/accommodation. Lastly the *listener’s standpoint,* as far as past musical experience, cultural affiliation and prior knowledge of the specific musical piece is concerned, plays a key role.

The scale’s two extreme positions (equilibrium- and perturbation-based anticipation) should be understood as theoretical ideals, while the area in between (historically based anticipation) is of a more universal character. The outer position of *equilibrium-based anticipation* involves the greatest degree of predictability. This position could also be called *future-listening,* as the anticipation is based on a certainty of the coming caused by prior veridical knowledge of the specific music piece that the listener projects into the future through protention. *Historically based anticipation* is to be understood as a somewhat omnipresent perception mode in the model according to which the listener is capable of performing informed guesses about the music’s further development on the basis of prior experience or *history*, including both veridical and schematic anticipation. This position could also be termed *past-listening* as the anticipation is based on musical structures in the past that stem from both the immediate past of the currently sounding musical piece through retention, as well as the listener's overall musical background comprising also whichever musical culture the listener may be a part of. *Perturbation-based anticipation* is the least predicting of the three listening positions. Here the listener is not familiar with the music and cannot either perform informed guesses about future events. This position could also be called *now-listening* as neither the musical past nor future can facilitate a qualified anticipation, which ideally leaves the listener unprepared in the mercy of a music happening right here and now.

The three anticipatory listening positions and their relation to HDM and VDM can be visually represented as in the figure below:

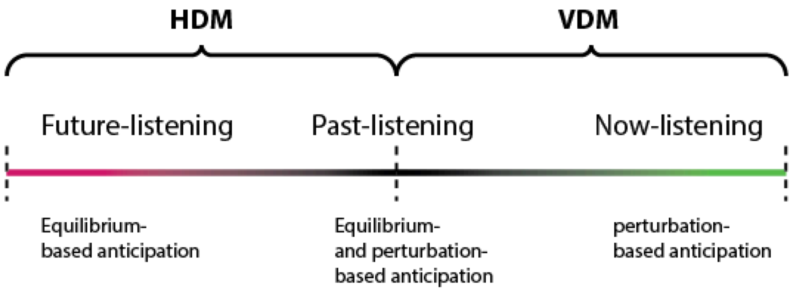


Figure 15 - Three anticipatory listening positions, HDM and VDM

The positions in the scale are expressions of theoretical and idealistic states and most often any given music will position itself somewhere between categories. One listening position does not exclude the other. It is highly likely that a listener during a piece of music will drift in and out between the three positions in a dynamic process. One moment the listener remembers precisely what is about to happen, in the next an informed guess is performed, which is either confirmed or denied, while later a situation could arise, in which the success rate of the predictions becomes close to zero. In addition, various simultaneously sounding layers of the same music can relate to different anticipatory listening positions. In genres such as blues and jazz, which are built on well-known musical structures but at the same time have a strong improvisational element, the listener can thus at the same time experience a predictable accompaniment and a partly or wholly unpredictable improvised solo.

The table below sums up the three anticipatory listening positions in terms of the aspects that may encourage them: Cognitive preconditions, listener's familiarity with the music, and musical qualities:

Equilibrium-based anticipation (future-listening)	**Historically based anticipation (past-listening)**	**Perturbation-based anticipation (now-listening)**
Stable cognitive equilibrium	Alternating between perturbation and equilibrium	Constant cognitive perturbation
Assimilation	Assimilation and accommodation	Accommodation
Listener is familiar with the music	Listener is unfamiliar with or partly familiar with the music	Listener is unfamiliar with the music
Music is very predictable due to: • Low musical entropy • Causal direction • Horizontal dominance • Much repetition	Music is partly predictable and allows the listener to make qualified guesses because the music: • Contains repetition • Belongs to recognizable style • Contains causal direction • Contains musical structures such as tonal language, harmonic progressions, melody etc. that are recognizable or easily learned	Music is very unpredictable because of: • High musical entropy • No causal direction • Vertical dominance • No repetition
Events take place through cyclic repetition (e.g. musical pulse, looping, repetitive chord progressions etc.)		Events take place through irreversible change

Table 1 - Three anticipatory listening positions with their respective cognitive preconditions; listener's familiarity with the music; and musical qualities

The table contains a number of defining aspects to the three positions. I will elaborate on these in the following pages. It is important at this stage to make clear that as the positions of equilibrium-based and perturbation-based anticipation are extremes or absolutes in the model, a musical listening situation may very well not include all of the mentioned defining aspects of a listening position and still be dominated by one position rather than the other. A piece of music may for example exhibit all other defining aspects of perturbation-based anticipatory listening except for an absence of repetition and still be highly associated with perturbation-based anticipation. Or a piece may contain a high degree of harmonic entropy, but will, by adhering to many of the defining aspects of equilibrium-based anticipation (such as a large amount of repetition, causally directed melodies, a clear pulse and easily comprehensible rhythmic patterns), nevertheless gravitate towards the equilibrium-based anticipatory listening position in the model. The defining aspects of the listening positions are thus not to be understood as categorical necessities but rather as contributors to the amount of gravitational pull each extreme in the model exposes the musical listening position to.

5.2.5. HISTORICALLY BASED ANTICIPATION

Historically based anticipation or *past-listening* is preconditioned by the listener being unfamiliar or partly familiar with the currently perceived music. The music may belong to a recognizable style or exhibits musical language, rhythm, harmony or other musical structure that is recognized or easily appropriated by the listener - thereby allowing qualified guesses about the music's further movement. In this anticipation mode expectations are constantly formed based on the newly experienced past in the listener's short-term memory as well as culturally conditioned cognitive schema of genre and musical tradition built up over time. It is also defining for this listening position that the actual further musical development when confronted by the listener's cognitive schema can lead to either *assimilation* of the perceived music or accommodation of the music-stylistic schema in that expectations are sometimes conformed, other times denied. Thus, an alternation occurs between equilibrium and perturbation, which in the discourse of Piaget's theory can be seen as an expression of learning. The listener learns the music, which in the process of past-listening shifts from being unknown to becoming known by the listener. As a result of this principle, the more times a listener is exposed to a piece of music, the more towards future-listening in the anticipation model the listening position will move. This is a point that will prove important when I turn my attention to the challenges of computer game music later in the thesis. Another characteristic of the listening position is that a "game" can take place between composer and listener in which music is designed to exploit and surprise in relation to expectations held in the listener's music-historic schemas.

It is typical for music that facilitates this listening position that it contains repetition and is based on musical structures such as harmonic progressions, melodies, gestures and so on that are recognizable or easily learned by the listener. Past-listening, as described here, is an expression of a way of listening to music, which is most often referred to when music is regarded from a phenomenological perspective, and the perceptual principle of constant dialogue between anticipation and the actual musical development must be assumed to be of such universal nature that it also exists in the model's two outer positions.

5.2.6. EQUILIBRIUM-BASED ANTICIPATION

Equilibrium-based anticipation or *future-listening* is dependent on the music being known by the listener to the extent that he or she can veridically recall precisely what will come in the immediate musical future - be it individual musical parameters such as melody, rhythm, harmonic progression and timbre or several parameters in combination. In some cases, the prediction may pose a close to complete impression of the present musical layers as a whole. Besides material from the listener's short-term memory and retention, the experience thus consists of precisely recalled musical events, progressions and structures located in long-term memory and projected into the future in the shape of an accurate protention of what is to come. This prediction exists on two levels. It builds on a precise recollection of the music and therefore a certainty about the further musical flow. This certainty allows the listener an imaginary pre-emption of what is coming before it occurs. At the same time, however, a process of historically based anticipation based on retentions create an expectation that in the state of *future-listening* is always confirmed. Thus, from this anticipatory listening position only assimilation takes place between the listener's cognitive schema and the actual musical development. There is within the frame of the listening experience complete congruence between the world and the listener's conception of the world - a stable cognitive equilibrium between the listener's expectations and the actual music.

A number of musical conditions act as conducive to a position in the model close to *future-listening*. The experience of low musical entropy, causal direction, horizontal dominance, much repetition and the presence of clear pulse or other musical events that occur through cyclic repetition are individually and collectively, due to their function as foundation for informed guessing, an expression of such conduciveness - although they cannot in and of themselves be said to fully guarantee cognitive equilibrium.

Examples of musical styles which are characterized by these qualities include some very simple country music, folk music and schlager music, but any music could in principle qualify for a position in the model close to this extreme position if it is heard repeatedly. Other styles that relate strictly to a particular tradition or conform to a particular standard, such as blues and some jazz will likewise have musical elements with a tendency to facilitate future listening, while other elements, as mentioned, can point towards other positions in the model. Music that promotes future-listening or equilibrium-based anticipation will often be phonographically recorded allowing repeated listenings and a deep and precise foreknowledge of the music. Moreover, consistency between the listener's music-cultural background and the musical tradition of the sounding music will also contribute to a given listening situation's gravitation towards this extreme in the model.

5.2.7. PERTURBATION-BASED ANTICIPATION

Perturbation-based anticipation or *now-listening* relies on the music being unknown to the listener. Music that encourages this listening position will typically be perceived as having: High musical entropy, absence of goal oriented causal direction, vertical dominance, absence of repetition, and will be characterized by musical events occurring irreversibly changing. Such music is ideally devoid of musical qualities, which could lead to informed guesses about the music's future. As with historically based anticipation and equilibrium-based anticipation, a constant dialogue occurs between the music of the now and the music of the past. But the listener will have difficulties assimilating the incoming music in the existing schemas, why these must be constantly accommodated. In the case of many accommodations of a cognitive schema the schema will with time be rejected as a model for the world and a new model will be constructed. In this way, according to theory, a stream of new schemas are constantly constructed, but again and again they must be accommodated due to incongruence between the actual musical sequence and the listener's expectations. These expectations are continuously denied, whereby the listener in theory will remain in a constant state of perturbation.

Examples of music that can potentially lead to this listening position, if the music is unknown to the listener that is, can be found within for example free improvisational music, aleatoric music, serial music and VDM. In addition, any music which belongs to a different music culture than that of the listener will likely have the tendency to push the listening position towards now-listening in a situation where the listener's culturally conditioned musical schemas do not contain the musical qualities constituting this thus alien music.

Cognitive schemas function on many simultaneous levels as a fundamental principle for understanding. It is important to mention that the experiences of the individual cannot be categorized into discrete and delineated schemas, but that the schemas should be seen rather as an immensely complex network of intertwined cognitive structures. Despite the efforts to break with tradition put into the ”new music” of early 20th century music, and attempts to devoid this music of the kind of *history* put forward by Ligeti, this ideally autonomous music does, however, eventually become history and thus, in Ligeti's terminology, a retrospective form the listener can base expectations upon. As such, it is possible for a schema to exist for the textural musical style represented by for example Ligeti's *”Atmosphêres”* despite its veridical unpredictability. Such a schema cannot be ignored in the current context, and *"Atmosphêres"* will in relation to this schema be assimilated. In this case, ”sound-mass music” as a schema is a predicate the listener puts on *”Atmosphêres”* and does not in itself preclude accommodation of other schemas on a music-perceptual level. Despite the predicate the music is still capable of functioning in the listener in for instance a gaming situation, just as ”romantic music” in a romantic movie, despite being potentially assimilated by the audience into the predicate ”romantic music” , may still function perceptually.

The question about which *listening level*[18] facilitates the experience is crucial here. If the player's attention is focused directly on the music, its function in the game - if the function is not *overtly trans-diegetic*[19] - would likely cease. It can be argued that the player in this situation is not immersed in the game. There is therefore a potential advantage in designing non-diegetic music to encourage what Truax calls *background listening* (Truax, 2001). More on this in chapter 7.

From the subject of three anticipatory listing positions I will now direct my attention towards another aspect of VDM perception and music perception in general, namely its ties to what has been called the *mirror neuron system*.

5.3. MUSIC AS EMPATHETIC CO-AGENCY

5.3.1. THE MIRROR NEURON SYSTEM AS A SHARED NEURO-FUNCTIONAL SUBSTRATE FOR MOTOR-ACTION, LANGUAGE, EMOTION AND MUSIC

A special class of neurons in the human brain called *mirror neurons*[20] are believed to behave identically to actions, whether the action is perceived or enacted by an individual. It is believed that such mirror neurons are responsible for the human ability for imitation and empathy by letting the observer intrinsically enact the observed - in effect practicing, for example, physical gestures as they are being perceived performed by others. In the following section a brief overview highlights some evidence pointing to a plausible link between such a system of neurons in humans and the experience of motor-action, phonetic gesture, language, emotions and music. This has relevance in the current chapter by allowing for a perspective on the experience of music as an intrinsic motor-functional, linguistic and emotional enactment of abstract musical form through a shared neuro-functional substrate, a human mirror neuron system. Such a perspective on the experience of music is included here to shed light on perceptual implications of the perturbation-based anticipatory listening position associated with VDM and to allow for the formulation of a list of perceptual attributes that may be associated to VDM.

---

[18] Barry Truax' notion of three *levels of listening* will be covered in chapter 7.

[19] Issues concerning the notion of *diegesis* are discussed in chapter 6 and 7.

[20] The mirror neuron system, although initially discovered in macaque monkeys, is widely believed to have a human equivalent in an area of the human brain called Broca's region. The reference to a "human mirror neuron system" made by e.g. Gallese (Gallese, 2008) is debatable. I am here referring to the equivalent human capabilities of Broca's region to the mirror neuron system found in monkeys as the "human mirror neuron system" for reasons of simplicity.

Mirror neurons were discovered in the premotor cortex of macaque monkeys through experiments including hand and mouth action - see for instance (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003). The behavior of these neurons was observed to be similar whether or not an action was observed by a monkey or carried out by it - in effect tying together our neurological understanding of perception and action. Similar neurological behavior in humans is described by, for example, (Grezes & Decety, 2001) and it is proposed that the monkey's premotor cortex, in which these mirror neurons were found, is homologous with an area in the human left frontal cortex referred to as Broca's area (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Gallese et al. propose that such a matching system in humans, similar to that of mirror neurons, could be involved in not only the recognition of actions but of *phonetic gestures* as well (Gallese et al., 1996). Based on theoretic work on speech perception by Liberman and Mattingly (Liberman & Mattingly, 1985); (Liberman & Mattingly, 1989), Gallese and co-authors describe phonetic gesture in the following way:

> ”…the objects of speech perception are not to be found in the sounds, but in the phonetic gesture of the speaker, represented in the brain as invariant motor commands. The phonetic gestures are 'the primitives that the mechanisms of speech production translate into actual articulatory movements, and they are also the primitives that the specialized mechanisms of speech perception recover from the signal' (Liberman and Mattingly, 1989).” (Gallese et al., 1996, pp. 607)[21]

Along the same lines Patrick N. Juslin's ”super-expressive voice theory” holds that musical expressions performed by voice-like instruments such as a violin may, by virtue of their resemblance to the human voice and by being able to surpass the voice's expressive capabilities in regards to, for example, speed and register, function as a form of musical exaggeration of speech (Juslin, 2001). Hereby the perception of emotional attributes normally associated with the human voice is projected onto the perception of music, which becomes ”*…a particularly potent source of emotional contagion.”* (Juslin & Västfjäll, 2008).

A link between the human equivalent to mirror neurons, first discovered in monkeys, and language is described by Gallese in his ‘‘neural exploitation hypothesis’’. Here it is proposed that mechanisms originally evolved for sensorimotor integration were later adapted as a shared neuro-functional substrate also for thought and language (Gallese, 2008). A communicative engagement of this substrate is supported by experiments with monkeys showing that the most effective visual stimuli for triggering mirror neurons were communicative mouth gestures like lip smacking, described in (Ferrari et al., 2003). It is believed that communicational

[21] (Liberman & Mattingly, 1989)

gestures such as language are particularly intimately linked to the mirror neuron system.

So what does this have to do with music? In regard to a connection between action, language and music through the mirror neuron system, Molnar-Szakacs and Overy hypothesize that the human mirror neuron system may mediate aspects of musical experience that are similar to the coupling between the perception and production of hierarchically organized sequential information known from language:

> ”…the powerful affective responses that can be provoked by apparently abstract musical sounds are supported by this human mirror neuron system, which may subserve similar computations during the processing of music, action and linguistic information.” (Molnar-Szakacs & Overy, 2006, pp. 235)

Furthermore, they suggest that language, music and motor-action may likely be similar in the way they recruit the mirror neuron system as a means of human communication through these modalities:

> ”…according to the simulation mechanism implemented by the human mirror neuron system, a similar or equivalent motor network is engaged by someone listening to singing/drumming as the motor network engaged by the actual singer/drummer; from the large-scale movements of different notes to the tiny, subtle movements of different timbres. This allows for co-representation of the musical experience, emerging out of the shared and temporally synchronous recruitment of similar neural mechanisms in the sender and the perceiver of the musical message. This shared musical representation has a similar potential for communication as shared language or action.” (Molnar-Szakacs & Overy, 2006, pp. 236)

It seems plausible that the perception and performance of music is tied to the perception and performance aspects of motor-action and language through this shared neuro-functional substrate.

Mirror neurons can additionally be linked to the intrinsic imitation of emotional states of others. They are proposed to be associated with high-level cognitive functions such as empathy (e.g. Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003) - that is, the ability to identify and experience not least emotional or affective states of other individuals. In neuroscience discourse this is referred to as *affective mimicry* (e.g. Ekman, 2014). And such capabilities for imitating emotions allowed by the mirror neuron system may likewise be coupled to music, as Molnar et al. suggest on the basis of neurobiological studies in emotional body language (Gelder, 2006)):

> “The expressive nature of any human action or vocalisation sends a signal of the intentional and emotional state of the executor, such that even footsteps can be correctly interpreted as conveying simple emotions

> (such as sad, happy, angry or stressed)." (Molnar-Szakacs & Overy, 2006, pp. 238)

It seems evident that the decoding of emotional information mediated through music mentioned above is equivalent to the kind of emotional information mediated through for example facial expressions. In other words, emotional and affective aspects of music perception can be seen here as linked to empathy. The receiver is able to empathize with the emotional state of the sender via a musical mediation, just as he or she can do so via mediations through language, facial expressions and other bodily gestures. This gives rise to very interesting questions in regard to computer generated music, which strictly speaking doesn't have a sender, as well as to the origin of motor-functional, linguistic, emotional and musical meaning when mediated through music. These questions are covered shortly.

Now, in regards to this link to the experience of music, the above hypotheses are based on an underlying linguistically anchored conception of music as an hierarchical organization of lower-level structures into higher-level structures as described by for instance Lerdahl and Jackendoff (Lerdahl & Jackendoff, 1983). Bharucha and co-authors propose an understanding of music as:

> "…a form of communication in which formal codes (acoustic patterns and their auditory representations) are employed to elicit a variety of conscious experiences." (Bharucha, Curtis, & Paroo, 2006)

In other words, the underlying conception of music here is defined by the music-structural attributes of *horizontal dominance* - that is, the musical characteristics I have associated with HDM.

5.3.2. MUSIC AS EMPATHETIC CO-AGENCY

The role of the human mirror neuron system as a shared neurological substrate to the action and perception of motor-action, phonetic gestures, voice-like musical expressivity, language, affective mimicry, and the experience of music brings me to suggest that the experience of music can be understood based on this shared substrate. In accordance with the above hierarchically defined music conception, but angled slightly differently, I will, for reasons that will be clear shortly, thus apply a definition to music as *that* which is experienced through a certain state of listening, *musical listening*, which in turn involves what I will refer to as "empathetic co-agency".

Empathetic co-agency refers to the above mentioned linguistic, motor-action, musical and emotional representation of the intentions of a sender as experienced by a receiver. I use the term *co-agency* here to address musical listening's coupling to the cognitive mappings associated with the discussed aspects of mirror neuron-founded embodied experience and language processing and to avoid the potential hierarchical connotation of terms such as *participation*, which may suggest a primary and a secondary agent in the activity in question. The notion of this co-agency as *empathetic* points to the involved emotional enactment. As such, the notion of *musical listening* as an expression of *empathetic co-agency* takes the consequence of the above discussed capabilities of the human mirror neuron system to link perception and action as well as to act as a shared substrate for language, motor-action, emotion and music.

The relationship between music as a form of communication (utilizing formal codes via musical units and their hierarchical organization at different musical time scales) and empathetic co-agency presupposes a degree of prediction, a degree of verification of protention. I am not able to enact the musical gesture if it is utterly incomprehensible to me, and therefore I am arguably not capable of relating to it empathetically. As mentioned earlier such successful protentions are connected with both future-listening and the omnipresent state of past-listening, but the further the listening position moves towards the extreme state of now-listening, the more the empathetic co-agency breaks down until, according to the above reasoning, the listener can no-longer *listen musically*. Crucially, I will argue that *musical listening* in this sense is associated with a certain flexibility in terms of the time period within which no musical patterns present themselves. One does not immediately fall out of a state of musical listening in patternless moments, it takes time - a period in which potential patterns are allowed to unfold and are sought constructed out of the perceived flow of sound. It is precisely this cognitive process of attempting to striate musically the flow of perceived sound in order to co-act empathetically with it that I refer to as musical listening. Such a time period may be prolonged or shortened due to circumstances of for example instrumentation and other forms of utilization of elements highly associated with music. A listener would arguably remain in a state of searching for musical patterns longer if the music is orchestrated for a symphony orchestra, which is schematically associated with musical listening, than if performed by a pot, an alarm clock and a bathtub.

Furthermore, it is from this perspective fair to hypothesize that a state of musical listening in these terms is associated with a sense of *control*. A sense of control that governs the musical listening state by coupling to the *agency* component of the *empathetic co-agency* notion. A sense of control is implied in the notion of *action,* and thus likewise in the notion of co-action. My point here is to argue for the likely correlation between the listener's ability to empathetically co-act with the music and the listener's sense of control.

This in turn renders the *vertical expression* (which was presented in chapter 2 and refers to the fusing of music-structural time and space into a *singularity* in which they can no longer be distinguished) *not* an expression of musical listening as it is per definition devoid of horizontal separable gestalts with which to co-act.

VDM, which plays with this singularity as an extreme gravitational pole, by positioning itself right at or very close to the border to this singularity, positions the listener at the fringes of complete control loss - only maintaing the listener in a state of musical listening by the degree of horizontality still present in the music and by being conveyed through means (such as a symphony orchestra as instrumentation) that keeps the listener in a state of looking for musical patterns. Within the singularity of the vertical expression music listening ceases and there is no longer a pursuit of gaining control over musical empathetic co-agency, therefore all perturbation, from a musical point of view, stops and the listening state shifts from being potentially stressful to resembling the sound of leaves on the trees, distant sound of the sea or traffic noise in which (like much VDM) many individual events constitute a mass of sound. In other words the music listening state associated with cognitive perturbation shifts to a non-musical listening state of cognitive equilibrium, and the vertical expression and its singularity between auditory time and space functions as a musical wormhole between the two extremes of the anticipation model[22]. In effect, the model of *figure 15*, which was presented previously in this chapter representing three anticipatory listening positions, can be bent to a circle. See *figure 16* below representing the circular model of three anticipatory listening positions, their relation to HDM and VDM as well as to musical listening, sense of control and the auditory singularity of the vertical expression:

[22] A movement into the auditory singularity may hypothetically also take place from the opposite direction. The musician Sting, when participating in the documentary, ”My Musical Brain” featuring an experiment in which his brain was scanned in an fMRI scanner while listening to a variety of musical styles, showed a remarkable change in neural activity when subjected to a piece of muzak. Compared to activity when listening to other musical styles Sting's brain all of a sudden showed no activity in the areas of the brain associated with listening to music (Pochmursky, 2009). Admittedly, the documentary cannot be used as a case of hard evidence to back up my case here, but it is, nonetheless, an interesting example of intriguing evidence to be followed up in future research that musical listening, as defined above, may cease, due to what might be interpreted as musical habituation and an extremely high degree of music-cognitive equilibrium.

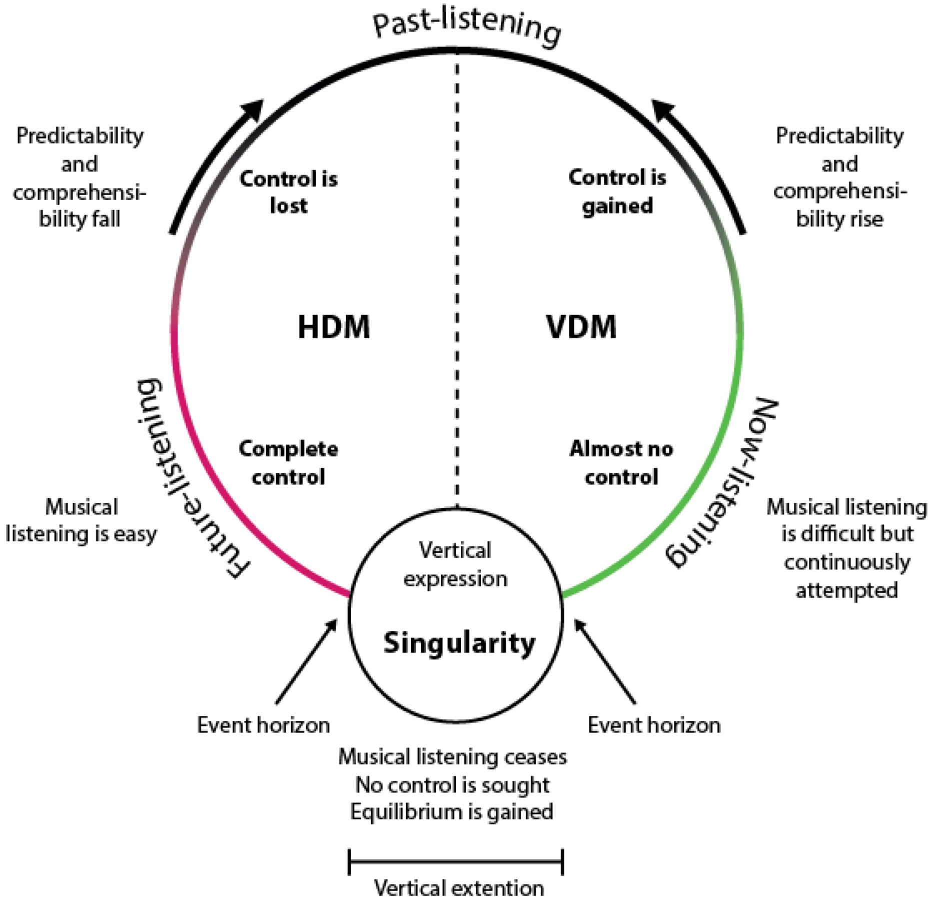


Figure 16 - Circular model of three anticipatory listening position

5.3.3. THE UNDISCLOSED "SENDER" OF VDM'S EMPATHETICALLY CO-ENACTED INTENTIONS

As mentioned in the previous section, the question of the *sender* from which motor-action, linguistic, emotional and musical intentions[23] originate is an interesting one, which I will go into here somewhat briefly in a philosophizing manner.

It can be argued that the actual existence of a human sender from which emotional information is mediated through music may in many cases be largely irrelevant to the experience of music. However, the perceived existence of a (perhaps illusory) sender, on the other hand, may be of some significance. It can be hypothesized that some aesthetically unsatisfactory cases of computer-generated music are perceived as unsatisfactory partly because of the absence of a clearly perceived sender - whether or not this sender is an illusion or not. It may likewise be hypothesized that some music, which *has* an actual human sender, is composed or performed in such a way that the presence of a sender is not perceived. (This would account for some of the difficulties one might have with enacting empathetically with cases of, for instance, perceptually highly entropic serial music.) This implies that from the perspective of the receiver, even an illusory sender of intentional information mediated through music may give rise to empathetic co-agency. On the other hand, it also implies that the presence of an actual sender does not guarantee empathetic co-agency.

I propose, however, that the concrete identification of a particular sender is not as significant to the musical listening experience as the intentions being received - whether these intentions are emotional or tied to language or motor-action. The sender may in this respect simply be regarded as nothing but the *music* itself. Highly interestingly in the context of this thesis, however, when the sender is not the performer or the composer, the perceived *origin* of such musically borne intentions is in question. This would be the case theoretically as a means for argument, but this question of origin is, I will suggest, at a perceptual level also an integral part of the musical listening state - not least as part of a multi-modal percept. This may pave the way for an understanding of the audiovisual couplings that take place on film and in games, between non-diegetic music and diegetic narrative elements.

I will go further into this now by opening slightly the notion of the *sender*. In a situation in which a monkey conveys meaning to another monkey through lip

[23] The notion of *intention* is here to be understood not in its mundane meaning as a form of motivation towards accomplishing this or that, but as a broader term encompassing mediated information about the state of the *sender* more generally. *Intension* as I use it here speaks particularly to the motor-action, linguistic and emotional aspects of the sender's state as represented through the communicatory capabilities offered by the mirror neuron system. My use of the term is thus related to, but not directly translatable to, the *intentionality* of Brentano or Husserl.

smacking, the relation between the sender (monkey A) the medium (visual and audible aspects of the smacking lips) and the receiver (monkey B) is clear. The same may be said when two humans are communicating through language and gestures. Musically, in a concert scenario, an instrumentalist may (simplistically speaking) convey intentional information through an instrument to an audience. But in many other contexts the sender might not necessarily be conceived of as a composer, singer or instrumentalist, nor as a human or a living conscious being altogether - whether actual or illusory. The question of the sender can, in music, be open. This is perhaps particularly true for acousmatic music, but also more generally it is not necessarily clear wherefrom the intentions being received through music are coming.

In audio-visual and interactive narrative contexts the notion of *sender* might to a large extent be attributed by the receiver to whichever narrative element of the multi-modal percept the music couples to. Such narrative elements would include the state of characters and avatars that can act as surrogates for a living sender whose state is decoded by the receiver, but, interestingly, they would also include the state of mystical phenomena such as ghosts, a presence of “something” that exceeds the boundaries of the senses or, indeed, of a narrative setting such as outer space[24].

Here one may ask a most crucial question:

> **What is the state of an undisclosed sender whose motor-functional, linguistic and emotional properties, when mediated through sound, produce VDM?**

I argue here that the perception of VDM (and of music in general) may be partly understood as the listener’s empathetically co-enacted ”answer” to this question - facilitating a musically borne but intrinsically co-enacted embodied realization of the state of this music’s disclosed or undisclosed sender. I furthermore suggest that such a realization, and the underlying intrinsic questioning of the origins of the intentions that led to it, may be seen as a foundation for audiovisual couplings between non-diegetic music and diegetic narrative elements - as part of the glue that ties the non-diegetic slow solo trumpet minor scale melody to the diegetic soldier and tells us he is lonely and in grief.

---

[24] In games, the trans-diegetic aspect makes the sender-medium-receiver relation difficult to uphold, as the trans-diegetic channel ties the communication to a loop - the player influences the narrative, which in turn influences the music, which in turn influences the player and so on. Such complications of the notion of diegesis in a game context is covered in the next chapter.

I shall turn to some implications of this observation now by formulating a list of perceptual attributes associable to this undisclosed sender in the context of VDM.

5.3.4. PERCEPTUAL ATTRIBUTES OF VDM - THE STATE OF THE SENDER

A list of qualities representing the intentional state of the sender of a VDM expression may at this point be formulated. These attributes can be summed up in the following adjectives: *Continuous, limitless, timeless, chaotic*[25]*, infinite, incomprehensible, uncontrollable, unknown, homeless* and *unpredictable*.

On the basis of the ground covered in this chapter I will approach this task from two angles: 1) based on the structural qualities of VDM discussed in the previous chapter, and their embodied perceptual implications as carriers of intention from an undisclosed sender and empathetically co-enacted by a receiver in a state of musical listening; and 2) based on these musical structures as they relate to the perceptual implications of perturbation-based anticipation.

In regard to the first angle, I will propose here that the adjectives of *continuous, limitless, timeless* and *chaotic* serve this purpose well. The continuous and limitless point to structural aspects of VDM discussed in the pervious chapter and deserve perhaps only little explanation here. This will follow shortly. The notion of *timelessness,* however, may give rise to a few connotations, why the meaning of the term, as I use it here, demands clarification. I am not referring to timelessness in the sense that a particular design may have qualities that are beyond fashion or are in some way valid in all periods throughout history. Furthermore, in regard to sound (as has been discussed at some detail in chapter 2), time is a necessity on a structural level. Physical sound is a manifestation of pressure variations in space over time. If one takes the time dimension out of the equation there can be no sound. So it is not timelessness in this manner that I am referring to, either. The timelessness I am interested in using as an adjective that may describe the perception of VDM is concerned with the mentioned perceptual absence of musical meter and the style's obscuring of horizontal gestalts. The term *vertical* is an ultimate expression of ”no time”. Musically, it represents a timelessness that comes to be, perceptually, by not allowing the subdivision of continuum mentioned in chapter 2, an avoidance of horizontal measurement facilitated by the covered structuring schemes of the style. This perceptual verticality, as already discussed in some detail, can be found in the singularity of the vertical expression. For VDM it is the position in the anticipatory model very close to this singularity that may give rise to a sense of timelessness.

---

[25] I.e. VDM can be seen to *hint* at the chaotic. It is itself not chaotic as such, but its breaking with the arborescent logic of HDM may seem fueled by some underlying force of chaos, of disorder.

As I went into in the previous chapter, VDM is highly associated with continuity, limitlessness, timelessness and the perceived chaotic in terms of its flirt with the horizontally unstriated qualities of the vertical expression and its function as a perceptual space-time singularity. On the vertical axis the expansions of vertical expressiveness covered in the previous chapter (such as an emphasis on mass-effect and the emancipation of pitch-space into spectral space as material for composition) also express a sense of continuity and limitlessness. Additionally, the three characteristics of VDM structure: *avoidance of horizontally separable gestalts, avoidance of perceivable horizontal irregularities* and a *weakening of the perceived bond between spectromorphologies* all contribute to a sense of continuity, limitlessness, timelessness and chaos conveyed through the music's resistance to horizontal - and in some cases also vertical - subdivision. VDM, seen as a musical percept that is embodied through empathetic co-agency, will, according to this trail of thought and by exhibiting these traits, project them onto an undisclosed sender.

In regard to the second angle, allow me to recapitulate the narrative focus of the thesis stated in the problem descriptions 2nd main objective:

> **…The focus is here on games that have as part of their narrative such phenomena as outer space, infinity, mystery, fear, divinity, paranormal activity - or generally, a presence of "something" that exceeds the boundaries of the human senses.**

It involves a list of narrative elements. These phenomena of: *outer space, infinity, mystery, fear, divinity, paranormal activity* or generally *a presence of "something" that exceeds the boundaries of the senses* might be summed up to express three primary narrative attributes: *infinitude*, *incomprehensibility*, *uncontrollability* and *the unknown. Infinitude* is here characterized as unlimited and unknowable; *incomprehensibility* as exceeding the capacity of the intellect; and the *unknown* is seen as an indefinite and un-manifested potential. Together these attributes express an *inaccessibility* on which *no control* may be exercised. In the following few paragraphs I will look into how each of the defined anticipatory listening positions relate to these attributes.

It is implied in the musical constituents of *equilibrium-based anticipation*: (low musical entropy, causal direction, horizontal dominance, much repetition and musical pulse) that they express finitude in the sense that they all contribute to a limitation of possible conclusions about the music's further development to only a few possibilities - ultimately to only one possibility, namely, the one known by the listener. Assimilation is preconditioned by experiences being comprehensible. In other words, comprehensibility is an expression of the ability of a schema to contain a given percept. A stable cognitive equilibrium governed by a constant congruence between percept and schema (the world and the individual's perception of the world) is also governed by constant comprehensibility, why *future-listening* is characterized by happening within the capacity of the intellect. Moreover, there is per definition no indefinite, un-manifest potential in the ideal state of stable cognitive equilibrium, just as there is no surprise because the listener is fully prepared for what is coming. Such a listening state is characterized by being experienced as completely accessible, and the musical listening taking place here implies a sense of complete control in terms of empathetic co-agency. Equilibrium-based anticipation or future listening, thus, categorically speaking, relates negatively in regard to having attributes in common with the phenomena mentioned above.

Regarding *historically based anticipation* and given the absolute and irreducible denotations of terms such as infinitude, incomprehensibility and the unknown it does not make sense to speak of a partial correlation between the music and these attributes. In cases where the listener during past-listening is taken either towards or away from experiencing the attributes as musical connotations, it will be as a consequence of attraction from the poles of the anticipation model rather than as an expression of qualities concerning historically based anticipation per se.

The musical traits of *perturbation-based anticipation* (high musical entropy, absence of causal direction, vertical domination, absence of repetition, and the fact that music events occur irreversibly changing) point towards an unpredictability whose ultimate consequence is now-listening. The unpredictability, these musical qualities promote, operates by virtue of the number of possible conclusions about the music's next steps being practically infinite. Due to unpredictability the listener is unprepared when faced with the flow of the music. The constant necessity of schema accommodation is an expression of the listener not being able to comprehend what is heard. Importantly, however, the very presence of the continuous construction of musical schemas is presupposed by what I have called *musical listening* previously in this chapter, and therefore it is dependent on the existence, to some minute degree, of patterns, horizontal separability between events or on being orchestrated or otherwise carried out by means that give musical associations, for any musical anticipation to be going on at all. The music is unknown and the concrete musical structures do not let themselves be immediately acquired because underlying musical structure does not let itself be realized or known. The anticipation may in particular be described as being directed against an un-manifested potential, something unknown, which in now-listening remains unknown, characterizing the music by an inaccessibility to the listener and implying a sense of total control loss in terms of musical empathetic co-agency. According to the reasoning in this and the previous section perturbation-based anticipation therefore welcomes the attributes of a presence of “something” that exceeds the boundaries of the senses.

As such, the further towards the extreme position of perturbation-based anticipation a listening situation places itself in the model the greater correspondence may be observed between the mentioned narrative attributes and the perceptual qualities of the music. In addition to the covered empathetically co-enacted relation to these attributes, the music thus functions also as a *metaphorical* expression of the attributes.

Furthermore, it can be said that VDM conveys *weightlessness* and *homelessness* due to the absence of a gravitational pull towards a tonal center on the vertical axis and metric subdivision on the horizontal axis. It is *timeless* in its refusal to striate the time dimension through rhythmic patterns and pulse. It is *limitless* and *uncontrollable* due to unpredictability and does not let itself confine to traditional musical rules.

This reluctance towards confinement within a set of perceived rules is a key feature of VDM perception, which I will go into in the following section.

5.3.5. MOVEMENT AT THE MUSICAL EVENT HORIZON

The metaphor of black holes and wormholes that I introduced in the introductory chapter of the thesis may at this point be expanded to encompass the concept of the *event horizon.* The event horizon is popularly speaking the edge of a black hole, a border beyond which nothing can escape the singularity at the center:

> ”The event horizon, the boundary of the region of space-time from which it is not possible to escape, acts rather like a one-way membrane around the black hole: objects, such as unwary astronauts, can fall through the event horizon into the black hole, but nothing can ever get out of the black hole through the event horizon. (Remember that the event horizon is the path in space-time of light that is trying to escape from the black hole, and nothing can travel faster than light.) […] Anything or anyone who falls through the event horizon will soon reach the region of infinite density and the end of time.” (Hawking, 1988, pp. 101)

I introduce this metaphor here, as it is the movement in and out of the auditory singularity, the crossing of the event horizon understood as a flirting with non-musical listening, that sets VDM apart from many other musical styles from a perceptual perspective.

VDM is governed by a set of underlying structural and perceptual principles. It is through the breaking out of the confinements of a musical system, through the violation of a schematic and veridical ”agreement” between the world and the subject - that the underlying forces of VDM come to effect. Some aspects of the structural characteristics of musical listening - such as perceived striation, patterns, the successful formation of gestalts or simply the use of musical instrumentation - is necessary in order to keep the listener in a state of searching for musical patterns - as a lure to facilitating musical listening, a frame to break out of. It is precisely in the motion of *breaking out,* in the process of *violation* that the underlying principles of VDM perception come into play.

The feeling you might get when hearing VDM in a horror scene may be hypothesized to correlate with fundamental existential fears of death; of dissolving into the unbound, unlimited, smooth universe transcending the limits of body; insanity as a disintegration of structured thought; a deeply rooted fear of chaos and utter control loss. The hypothesis being that such fears are potentially promoted for example in situations where our conception of the world is realized to be false - leaving us defenseless to its whims. As such they might be associated with an encounter with *continuum* - an encounter which suggests to us a cancellation of our ability to group and segregate our stimuli into meaningful and manageable gestalts. Furthermore, from the perspective of embodied experience, the sense of control loss may be associated with also a sense of bodily control loss. As an expression of empathetic co-agency, the listener may be led to co-enact an abstract movement that disregards or transgresses the boundaries of the human body and mind.

But vertical dominance's defiance of horizontal striation may lead to musical experiences not only of negative valence but also of positive valence depending on context. The potential scary character of VDM does not come from the vertical dominance alone, although a scary character may very well be supported by vertical dominance. Any threatening expression resides primarily in the organization of the vertical expression and its horizontal development - and these aspects of VDM follow similar principles of perceiver *competence*[26] as HDM. It is tempting to regard VDM as an inherently frightening musical style because the musical qualities associated with some works of its pioneers such as Penderecki and Ligeti are stereotypically regarded as relatively unpleasant to listen to. This is supported by these composers' music being affiliated with not only science fiction and mystery but also with horror films[27] as also mentioned in chapter 4.

The relationship between *vertical dominance* and *vertical expressions* is perhaps best understood by seeing the vertical expression's emotional attributes as the primary source for affect and secondly the vertical dominance as setting this source for affect free of its horizontal constraints of patterns. This would imply that potentially frightening vertical expressions (stereotypically exhibiting dissonant harmony such as a chromatic cluster, a harsh timbre and erratic textures like very arhythmical string tremolos) become frightening in a different way by the added element of vertical dominance with its unique associated perceptual attributes, intentions and potential for conveying that which is *continuous, limitless, uncontrollable, chaotic and timeless*. On the other hand a friendlier vertical expression (which according to typical musical clichés could be comprised of e.g. consonant harmonies, soft timbres and a smooth texture of gently bowed strings),

[26] The notion of *competence* is covered in the next chapter.

[27] Some of Penderecki's works associated with film were mentioned in the previous chapter. Examples of Ligeti's VDM on film include: Stanley Kubrick's ”2001: A Space Odyssey” from 1968 featuring ”Atmosphêres”, ”Lux Aeterna”, ”Aventures” and ”Requiem”; Peter Hyams' 1984 follow-up ”2010: The year We make contact” featuring ”Lux Aeterna”, Kubrick's horror classic ”the Shining” from 1980, which features ”Lontano”; Scorsese's psychological thriller ”Shutter Island” also using ”Lontano”; and Gareth Edwards' monster science fiction movie ”Godzilla” from 2014 in which Ligeti's ”Requiem” is featured.

when supplemented with vertical dominance, sets free, so to speak, then *these* musical qualities. As such this latter example becomes a synthesis of generally pleasant sounding vertical expressions with the continuous, limitless, uncontrolled, chaotic and timeless - a musical expression entirely different, and with perhaps more potential of promoting associations to phenomena of positive valence like divinity, bliss, deep relaxation and meditative states of consciousness.

5.4. CONCLUSION

The model, ”three anticipatory listening positions” is comprised of *equilibrium-based anticipation* (or future-listening), *historically based anticipation* (or past-listening) and *perturbation-based anticipation* (or now-listening). These positions are tied to the different musical structuring schemes associated with HDM and VDM. VDM, in particular, is connected with the extreme position, *perturbation-based anticipation*, by tending to cause a constant cognitive perturbation and by forcing the listener to constantly adapt his or her schema of the music through accommodation rather than assimilation. This process ideally requires that the listener is unfamiliar with the music, which may be very unpredictable by exhibiting high musical entropy, no causal direction, an absence of repetition and, finally, by vertical dominance (therein implied the characteristics of vertical domination, covered in the previous chapter, of avoiding horizontally separable gestalts, avoiding perceivable horizontal regularities and weakening the perceived bond between spectromorphologies).

Other perceptual attributes of VDM include a potential feeling of control loss which is connected to VDM’s special quality of moving in and out of *musical listening* and thus playing with the listener’ ability to *empathetically co-enact* the musical structures as they happen.

Thus the *vertical expression* and its quality as a *singularity* is not music. What happens when perceiving a purely vertical expression (not VDM, but the vertical expression by which VDM is dominated) is not musical listening. This is argued with no intention of implying any sort of underlying devaluation of the potential beauty one might encounter in such an experience and with no agenda to suppress the significance of reduced listening[28] as a channel for auditory satisfaction. Without going too deeply into the complex (and old) question of what ”music” is, I take the position that music regarded as ”a state of listening” holds strong potential for shedding light on the perceptual workings of VDM in the context of the present research. On such grounds, ”music” can be regarded as *that* which is experienced when in a state of *musical listening*.

[28] Schaeffer’s term for listening to the sound itself as opposed to e.g. identifying its sound source.

Musical listening as *empathetic co-agency* describes a participatory action in which the listener, through empathetic capabilities tied to the mirror neuron system, enacts the music while listening. This in turn requires an ability (at least to some extent) to predict the flow of the music. I have hypothesized that this ability is only required to some extent because musical listening might be satisfactorily occurring through an attempted identification of musical patterns at the meso-time scale - as such it is less significant if any musical patterns are found as long as they are *attempted* to be found. Empathetic co-agency, which as a term combines linguistic, motor-functional, musical and emotional aspects of mimicry allowed by the human mirror neuron system, is dependent on various degrees of hierarchical structures, patterns and musical horizontality or other means with clear musical associations. I have argued in the present chapter that both on a music-cognitive level as well as at the level of embodied experience, a direction in the model towards perturbation-based anticipation is associated with a sense of control loss, whereas the direction towards equilibrium-based anticipation lets the listener gain a sense of control over the empathetic co-action.

I have addressed the question of an undisclosed, unidentified, non-concrete and perhaps illusory *sender*, who takes the role as an origin for *intentions* within a linear sender-medium-receiver communicational paradigm. Such a linear paradigm is problematic in the context of games. This issue, which is taken up in the next chapter, does not, however, cancel the validity of the argument in regard to VDM perception on a general note. The notion of a sender in this respect was seen as occupied with a certain *state* constituted by *intentions* that are communicated through music and with which the receiver co-enacts empathetically. This empathetic co-agency with a sender was, I have suggested, intrinsically present in the musical listening state. This led to a crucial question in regard to VDM perception:

> **What is the state of an undisclosed sender whose motor-functional, linguistic and emotional properties, when mediated through sound, produce VDM?**

I proposed that the perception of VDM may be partly understood as the perceiver's empathetically co-enacted "answer" to this question - in the form of a mental representation based on mimicking capabilities of the mirror neuron system that facilitates a realization of the sender's state through music listening. I further argued that this role of a sender of intentions may be attributed by the perceiver to narrative elements of a multi-modal percept - thus audio-visually linking non-diegetic music and diegetic narrative elements like characters, avatars, the presence of "something" that exceeds the boundaries of the senses, or the narrative setting such as, for example, outer space.

A list of adjectives was presented as descriptors for perceptual attributes associable with VDM. This association was based on perceptual attributes of perturbation-based anticipation as well as the perspective on musical listening as empathetic co-agency. These adjectives are: *continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless* and *unpredictable*.

Finally, I argued that emotional attributes of VDM may be of positive or negative valence, and that it is primarily the qualities and horizontal development of the vertical expression that is responsible for the polarity of this valence. Crucially, however, the perceptual attributes of vertical dominance play a highly significant secondary role, which changes the quality of such emotional attributes of the vertical expression by adding its own unique qualities tied to said adjectives.

Questions concerning the music's ability to fuse with for example narrative elements such as outer space in multi-modal media such as film or computer games and thereby giving outer space a voice, making it an audiovisual expression rather than a purely visual and narrative phenomenon, will be taken up in the next chapter. Here VDM's broader potentials for narration in such contexts are examined.

CHAPTER 6. NARRATIVE POTENTIAL OF VDM

From mapping out a set of perceptual implications of VDM it is now time to take a look at how such implications are involved in VDM performing a narrating function. The following study of VDM's narrative potential not only adds semiotic aspects of VDM to such perceptual implications of narration but also directs attention towards outer space as a narrative setting.

6.1. INTRODUCTION

Understanding the function of non-diegetic music as a co-constituent of a narrative is important in the context of this thesis - not least in regard to its aim to understand the impact it would have on the gaming experience to implement VDM in a game. Furthermore, an investigation of VDM's narrative potential is a key component in the effort to create a structural, aesthetic and functional foundation from which a generative VDM system may be later built. The ability of music in general, as well as VDM specifically, to narrate in films and games, and the way in which this narrational potential is manifested, can be a complex and ambiguous subject, and I will focus here on only a few major aspects. In regard to the question of *how* and *what* VDM may narrate in a multi-modal context such as in computer games, I will discuss four angles.

- The role of different *diegetic layers* in the formation of multi-modal meaning
- Competence-related *semiotic attributes* of VDM
- Empathetically enacted *perceptual attributes* of VDM
- Four aspects of VDM's *relation to its multi-modal context:* comprising *general* and *specific* relations that can be based on *game-internal* and *game external competences.*

Together these four angles give insight into the processes that let VDM, and music in general, take part in the narrative in multi-modal and interactive contexts.

The first angle deals with the role of different *diegetic* layers in the formation of multi-modal meaning. Initially, this demands a discussion of the applicability of the term, *diegetic*, in regard to games, as this can be problematic for several reasons. Additionally, the idea of trans-diegetic music, which through the player may have an influence on the diegesis, poses some issues in regard to the linear sender-music-receiver communicational paradigm, which was set up in the previous chapter as a foundation for seeing music as empathetic co-agency. I argue in this chapter that the communicational paradigm involved in experiencing game music may be either linear or circular. And that, from the point of view of the player, a linear communicational paradigm is acting when trans-diegetic functions happen *covertly*, while a circular model relates to trans-diegetic functionality being *overt*. As a final concern in regards to diegesis, I argue that the subtle mystery associated with the very presence of a non-diegetic layer may help to support the ”undisclosedness” of the undisclosed *sender* whose state through musical mediation is realized by the player on a narrative meta-level - a level that is fundamentally different from that offered by diegetic sound. I call this subtle mystery the ”great beyond of the non-diegetic”.

Secondly, the notion of *competence* is discussed as a model for explaining parts of the narrative potential of VDM arguing that part of this potential is fulfilled through *competence-related semiotic attributes* of the style. From a semiotic perspective VDM, as a signifier, can be seen to have acquired a set of connotations that to various extents is present in the competence of the players of games and audiences to films that contain VDM. This part of the chapter sees VDM as a semiotically charged musical sign whose meaning is likely decoded by an audience via a relatively specific media competence.

The third angle recapitulates the findings of the previous chapter and finds that VDM may narrate also through the conveying of intentions of an undisclosed sender whose state is mediated through music. I argue that this state is projected onto elements of the diegesis. In effect, principles of *musical listening*, as defined in the previous chapter, enables VDM to fulfill part of its narrative potential through *empathetically enacted perceptual attributes* of the style.

The fourth angle argues that competence-related semiotic attributes and empathetically enacted perceptual attributes may relate to a game on either a *general* or *specific* level - to overall game characteristics and aesthetics or to narrative subelements respectively. I will argue that when narratively coupling *specific* vertical substructures of VDM to *specific* in-game elements, a degree of vertical separability is required, and that the auditory gestalt principles, highlighted by Bregman and McAdams in their theory of auditory stream analysis, may be of use for this purpose. I go on to arguing that the semiotic competence of players may be based *game-externally* or *game-internally* where decoding is conditioned by either schematic or veridical memory respectively.

In the next section of the chapter, called ”Outer space as narrative setting”, I put my focus in the direction of *outer space* and this narrative setting as a component of player competence. Outer space here acts as a concrete example of a cross-modal narrative coupling to VDM, which may suit the style’s semiotic and perceptual attributes well. In this section, an overview of sound in the golden age of space oriented science fiction is provided together with a discussion of the use of infinite reverberation decays as an expression of some of the same perceptual attributes as are associable to VDM. The sections conclude by arguing that VDM shares characteristics with early sounds of the space oriented science fiction genre through musical mediation of movements in an *inner* space as a base for humans' realization of the otherwise impossible to grasp *outer* space.

6.2. DIEGESIS

6.2.1. APPLICABILITY OF THE TERM IN REGARDS TO GAMES

The concept of *diegesis*[29] (Loosely translatable to ”narrative” from Greek) is often used in audio-visual theory in the context of film. Diegetic events are described by Bordwell & Thompson as: *”events taking place in the story world”* (Bordwell, Thompson, & Ashton, 1997, pp. 331). The question of diegesis is relevant in the current chapter because different diegetic layers (diegetic, non-diegetic as well as overtly and covertly trans-diegetic) may be seen to have different narrative potentials. In particular, music that is experienced as non-diegetic, I will argue in the following pages, may project its semiotic and perceptual attributes cross-modally to other elements of the game due to this non-diegetic status.

Diegetic sounds are experienced as belonging to the fictive world of the narrative. Non-diegetic sound on the other hand is experienced as coming from a source outside of this world. A traditional film score, whose sound source does not originate within the fictive narrative world, is thus referred to as non-diegetic, while the dialog, foley and sound effects are typically regarded as diegetic (e.g., Metz, 1974). However, it is problematic to uncritically transfer the concepts to the computer game media, since they inter alia do not take into account the impact of adaptive sound and music as catalysts for the player's decision-making processes. An adaptive music that otherwise occupies a role in the game similar to that of a non-diegetic film score, can via the player influence the course of action as he or she reacts to the music - effectively erasing the boundary between the dietetic and non-diegetic narrative layers and arguably rendering the terms meaningless in interactive contexts. Kristine Jørgensen introduces the term *trans-diegetic* to account for such a scenario. A given sound receives a trans-diegetic function by *”...transcending the border between diegetic and non-diegetic* (Jørgensen, 2008, pp. 85, translated from Norwegian). The use of the diegesis concepts is, however, still problematic in the context of computer games according to Jørgensen - even after the addition of trans-diegesis because this new concept is based on a dichotomy that from the above reasoning cannot be unambiguously defended in the presence of adaptive music.

[29] Before Bordwell and Thompson, Christian Metz used the term in relation to film. Plato speaks of *diegesis* in ”The Republic”, Aristotle in ”Poetics”. The term has an opposition in the term *mimesis,* which is associated with mimicry and en-acting as opposed to the narrating associated with diegesis. One could argue in this respect that in regard to interactive media such as computer games, the term *mimesis* is in fact highly applicable.

Jørgensen later leaves behind the idea of trans-diegesis as a useful solution to the problem in favor of regarding computer games as *interfaces* (Jørgensen, 2010). Espen Aarseth distinguishes between the computer game's *gameworld* and the linear fictional world, by which the concept of diegesis is defined. Emphasis in this connection is put on the computer game reality's key attribute as a functional and playable gameworld or arena (Aarseth, 2008, pp. 118). They hereby reject categorically diegesis as a useful concept in a computer game context.

Although I must agree that it is problematic to equate the linear fictional world of film and the interactive gameworld, and concede that it is difficult to maintain a contradictory relation between diegetic and non-diegetic in computer games, I argue that there are equatable relations between the cinematic narrative *imaginary world*[30], which is experienced by an audience, and the interactive *imaginary world* in which the player is immersed, that will make the concept nevertheless highly justifiable. Despite some very significant differences between the two media, not least rooted in computer games' functionality and interactivity, there are some equally significant - and perhaps quite obvious - similarities. Firstly, there is a basic similarity in that both imaginary worlds by virtue of their immersive qualities are experienced as audio-visually presented alternatives to a reality outside the medium. In both cases this alternative is presented visually on a screen[31] and audibly through speakers. In continuation of this, because of the huge diversity of different interactive programs, we call computer games - each with very different characteristics - one cannot ignore the fact that some games bear such a great expressive similarity to movies that part of our prior experience - or *competence*[32] - as a movie audience must be assumed to follow us into the gaming experience. Visual and auditory realism; the presence of characters; a basic plot or narrative frame in which events take place; as well as some games' use of cut-scenes (cinematics), which more or less seamlessly lead the player in and out of interactive and non-interactive scenarios, are examples that represent such expressive similarities and may promote competence projection between media. There does exist, therefore, from a phenomenological perspective, the possibility of considering certain computer games as a kind of "movie you can play in." This would apply to games such as *"BioShock"* (2K_Games, 2004) or *"Halo"* (Bungie, 2001) and generally to most First Person Shooter type games with an elaborate storyline. As part of a media culture in which both films and computer games appear, the vast majority of players are accustomed to experiencing the non-diegetic music associated with film. In this regard, it is fair to assume a habituation

[30] *Imaginary world* understood here as a phenomenological, intrinsic pendant to the extrinsic narrative world or gameworld.

[31] Virtual reality and the use of goggles such as the Oculus Rift may be seen to have pits won unique implications in regard to diegesis. I will not, however, go into such aspects in this thesis.

[32] The notion of *competence* is covered later in this chapter.

to the non-diegetic narrative layer which results in an expectation amongst players of the presence of such a layer - thereby causing him or her to position computer game music, whether it is adaptive or linear, in the same ”perceptual slot” as non-diegetic film music.

I dispute, therefore, a categorical rejection of the concept because such a rejection, in the effort to define computer games as a sovereign independent media, denies any significance of habituation to the non-diegetic layer as a perceptual phenomenon that can be projected by the player from movie to computer game. As a consequence of this phenomenological approach to the concepts of diegetic, non-diegetic and trans-diegetic, I will use them in regard to computer game music, not in spite of, but exactly because of their association with the film medium.

6.2.2. THE QUESTION OF THE *SENDER* IN OVERTLY AND COVERTLY TRANS-DIEGETIC MUSIC

In interactive contexts, such as games in which music may be trans-diegetic, the sender-medium-receiver communicational paradigm discussed in the previous chapter is to some extent no-longer valid. The interactive dynamic between the player, adaptive music and gameplay can be seen as an expression of a non-linear and circular system rather than a linear one.

So what happens with the notion of the *sender* when the communicational paradigm is no-longer linear but circular? From the perspective of a projected media competence it can be argued that nothing much changes. The music retains its occupancy of the projected non-diegetic narrative layer. From here it may couple to diegetic narrative elements - partly due the illusory presence of a fully undisclosed sender, whose state is mediated musically through empathetically co-enacted motor-functional, linguistic and emotional intentions.

However, if the music’s adaptive functionality is obvious, the question of the sender is partly answered as - the *player*. This answer is likely to occur if the player can easily see through the rules that facilitate the interaction between his or her actions and the musical response to these actions. (If, for example, every time the player enters a room in the game, the music plays a certain phrase). In other words, *direct adaptive music* (Collins’ notion of *interactive* music) is more likely to be figured out in this way than *indirect adaptive* music (Collins’ *adaptive* music). This scenario makes the question of the sender (a question which I argued in the previous chapter, can be seen as intrinsically present in the musical listening state) less of a mystery and makes the music subject to conscious manipulation by the player in an obvious way[33]. Based on the reasoning put forth in the previous chapter about the undisclosed sender as a means for cross-modal couplings, such a case of adaptive

[33] Games such as Guitar Hero (Harmonix, 2005) or Rockband (Harmonix, 2007) can be considered extreme examples in which the interactive functionality is very obvious.

music, in which the interaction is obvious, can be expected to weaken the cross-modal link between the music and the diegetic narrative element it is intrinsically coupled to by the player. Here one may speak of the trans-diegetic function of the music being actually *perceived* and the problems posed by, for example, Jørgensen (in regard to the use of the terms diegetic and non-diegetic in a game context) become somewhat relevant also from a phenomenological perspective. This leads me to propose that in terms of the *sender,* two communicational paradigms exist for adaptive game music.

Firstly, a *linear* sender-music-receiver paradigm exists when the music retains its occupancy of the non-diegetic narrative layer due to the projection of this layer by the player, based on prior experience with the film medium. A game's resemblance to a "movie you can play in" further promotes the existence of this paradigm - the linearity of which is phenomenologically valid when any trans-diegetic functionality the music might have is *covert*. See figure below:

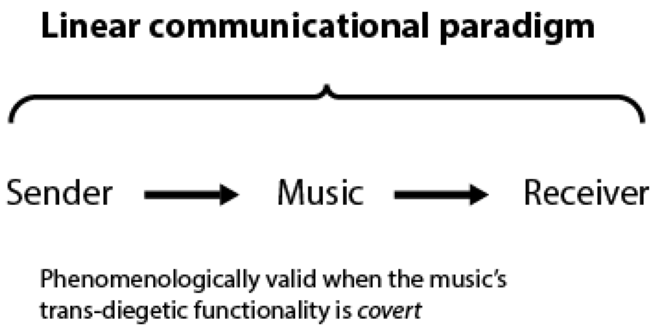


Figure 17 - Linear communcational paradigm

Secondly, a non-linear, circular communicational paradigm may exist between the sender, the receiver (player) and the music. This paradigm is phenomenologically valid when the music's trans-diegetic functionality is *overt*. See figure below:

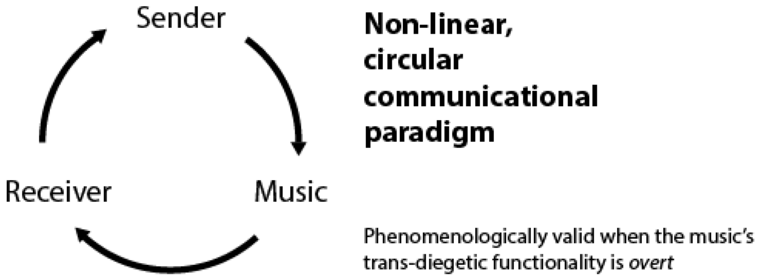


Figure 18 - Circular communicational paradigm

Importantly, even in the case of the latter paradigm, the player only ever to a limited degree takes the role of sender. Although the player has an obvious influence on some aspects of the music (the aspects which have been designed to be obviously adaptive) significant characteristics of it remain un-manipulatable to the player. Thus these remaining untouched characteristics of the music still convey information about the state of a (partly) undisclosed sender through musically mediated and empathetically co-enacted intentions. The music's ability to convey such information, as well as its ability to couple to elements of the diegesis, will thus vary depending on the extent to which the music is obviously manipulatable. Consequently it can be said that the more blatantly obvious the influence of the player's actions are on the adaptive music, the weaker is the music's coupling to diegetic narrative elements and the weaker is the conveyed information about the *state* of such elements. This makes the upholding of a border between the obviously trans-diegetic and the phenomenologically non-diegetic narrative layers an important issue in regard to adaptive game music's narrative potentials.

The distinction between overt and covert trans-diegetic functionality, which I am proposing here, is useful in order to address, for instance, the practical problem that may occur when the adaptive music system reveals itself. If overtly trans-diegetic, it may become predictable what the music will do when certain events occur in the game. The BAFTA award winning composer team behind the music for the game *"Alien: Isolation"* (Assembly, 2014) referred to problems of this kind at the Game Music Connect conference in London 2015. In their work with implementing adaptive music in the game they had to be very careful not to reveal that an alien was about to attack the player's avatar, and so give away the surprise (Christian Henson & Smith, 2015). However, in other games, such music games like Guitar Hero and Rockband, the overt trans-diegetic function of the direct adaptivity of the music may serve as a deliberate and important game-play component. In both examples the notions of overt and covert trans-diegetic music are of direct consequence to both the gameplay and the quality of the gaming experience.

6.2.3. THE *GREAT BEYOND* OF THE NON-DIEGETIC

The existence of a clear border between the phenomenologically diegetic and non-diegetic is likewise of significance to the way in which music may narrate in games. This is not least because the non-diegetic narrative layer can be seen to contain special narrative potentials of its own. Setting aside for a moment the audience' and players' preconditioning to experiencing the coexistence of diegetic and non-diegetic layers, if one is to see the mediated world of film or games as analog to the non-mediated reality it often seeks to emulate, there is something inherently mysterious about the existence of a non-diegetic layer. Its existence poses unanswered questions. What is it? Where does it come from? What does it represent? By its function as separate from the story world's diegetic reality, non-diegetic music suggests, I argue, a ”great beyond”, which, as media competence grows, becomes a phenomenologically preconditioned imaginary space potential. The non-diegetic layer can be seen as an abstract field of latency occupied by some unknown entity whose *state* is only realizable through the more or less ambiguous qualities of the music inhabiting it. This aspect of the non-diegetic layer supports the argument for maintaining a distinction between diegetic and non-diegetic, also in games, because a *great beyond* implies a border between two different layers of the narrative.

I will regard the subtle mystery posed by the presence of this narrative layer as supportive of the ”undisclosedness” of the sender in a multi-modally anchored linear sender-music-receiver communicational paradigm. From this perspective, the vacuum created by the unanswered questions posed by the presence of non-diegetic music facilitates this music's functionality as a means for empathetic co-agency by allowing it to exist phenomenologically free of the diegesis. (An important issue as discussed above in relation to overt and covert trans-diegesis). Along the same line of thought as the discussion of the undisclosed sender in the previous chapter, the non-diegetic layer hereby offers a ”space*”* of *undisclosed origin* for the music to occupy. This origin can be hypothesized, like the musically mediated state of a *sender*, to be intrinsically ”answered” by the receiver, who in the process synthesizes a realization of the multi-modal construct as a whole.

6.3. VDM AND MULTI-MODAL PERCEPTION

The notion of *multi-modal perception* is among others theorized by Birger Langkjær in relation to audio-visual fiction. Here it is seen as an expression of a process by which meaning is created by the observer on the basis of stimuli. These stimuli are separate on the technical level as sound and image, separate on the sensory level as hearing and seeing, but unify to one meaning on the semantic level (Langkjær, 2000, pp. 24). According to this premise, any given music which is implemented in a film will cause the audience to relate the music's semantic attributes to other elements of the media - just as attributes of these other elements will be projected onto the music. The audience thus creates one coherent synthesized meaning out of the multi-modal phenomenon. In computer games, as discussed above, the situation is a bit more complex because there is no stable sender-receiver relationship between the player and the medium due to the music's and other game sounds' often trans-diegetic function. Nevertheless, there seems to be nothing to suggest that music's performance of trans-diegetic functions disarms its ability to attribute to other game elements' semantic, semiotic or empathetically co-enacted meaning. In other words, through principles of multi-modal perception music holds a narrative and aesthetic expressive potential of inducing other media elements with its meaning.

6.3.1. MEDIA COMPETENCE AND VDM'S POTENTIAL FOR NARRATION THROUGH SEMIOTIC ATTRIBUTES

The narrative potential of VDM is shaped also by the semiotic meaning it itself has established through years of use on film and, more recently, in games. There are certain attributes now associated with VDM by film audiences and gamers, who, through more or less consistent exposure to this type of music in specific filmic contexts, have learned its symbolism. Film audiences and gamers have acquired a certain media *competence.* Anahid Kassabian describes the notion of competence in regard to film as follows:

> ”The skill that generates consistency in encodings and decodings of film music is “competence.” Clearly, competence in this sense can only function for speakers (and listeners) of the same language (or musical genre), and the consistency will vary according to fluency (extent of experience in the genre), personal history, etc. Competence is a culturally acquired skill possessed to varying degrees in varying genres by all hearing people in a given culture.” (Kassabian, 2001, pp. 20)

The question of how VDM may be said to narrate in computer games can, as I have already discussed, be partly answered in the light of this notion of competence. By triggering associations and semiotic meaning contained in the player's competence of VDM in film and games, the style mediates a set of competence-related semiotic attributes by which VDM may narrate in a multi-modal context. These semiotic attributes are, to at least some extent, constituted by VDM's association with horrific and mysterious scenarios on film and in games through the years, including as a non-diegetic companion to scenes situated in outer space.

As a result of an audience's competence with associating VDM with horror and mystery it is more than likely that the style will be decoded along these narrative lines when used in new productions. Thus VDM, as a matter of competence, can be seen to support such a narrative focus in any new games and films of which it is a part.

6.3.2. THE UNDISCLOSED SENDER AND VDM'S POTENTIAL FOR NARRATION THROUGH PERCEPTUAL ATTRIBUTES

From a perspective where music is defined as *that* which is listened to in a state of *musical listening,* and where musical listening is conceived as a process of *empathetic co-agency,* the previous chapter led me to ask: *What is the state of an undisclosed sender whose motor-functional, linguistic and emotional properties, when mediated through sound, produce VDM?* I argued that the perception of VDM may be partly understood as the listener's empathetically co-enacted ”answer” to this question - facilitating a musically borne but intrinsically co-enacted embodied realization of the state of this music's disclosed or undisclosed sender. I further suggested that such a realization - as well as the underlying intrinsic questioning of the origins of the motor-functional, linguistic and emotional intentions that led to it - may be seen as a foundation for audiovisual couplings between non-diegetic music and diegetic narrative elements.

From this point of departure, I formulated a list of adjectives to represent the qualities of such a sender's intentional state: *Continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless* and *unpredictable*. The list of adjectives was devised by taking into account the structural properties of VDM[34] discussed in chapter 4 as well as the perceptual implications of perturbation-based anticipation[35] discussed in chapter 5.

I will consider these findings an account of VDM's empathetically co-enacted *perceptual attributes.* Thus, in addition to the semiotic attributes recently discussed, VDM may also narrate by virtue of the cross-modal coupling of this set of perceptual attributes to other components of the multi-modal phenomenon.

6.3.3. VDM'S GENERAL AND SPECIFIC RELATIONSHIP TO THE MULTI-MODAL CONTEXT

The relation of the semiotic and perceptual attributes of VDM (and music in general) to other components of a multi-modal construct that the music co-constitutes can be said to occur in at least two ways:

- through the coupling of VDM's *general* semiotic and perceptual attributes to *general* aesthetic and overall narrative directions of the multi-modal phenomenon

[34] I.e. comprising three prominent characteristics of VDM: avoidance of horizontally separable gestalts, avoidance of perceivable horizontal patterns and a weakening of the perceived bond between spectromorphologies.

[35] I.e. cognitive perturbation due to high musical entropy, absence of causal direction, vertical domination, absence of repetition, and the fact that musical events occur irreversibly changing.

- through the coupling of VDM's *specific* music-structural constituents to *specific* components in other modalities of the multi-modal phenomenon

A notable presence of VDM in a game or film induces in itself a *general* kind of narrative and aesthetic meaning to the overall game or filmic expression. Through this general relationship, the competence-related semiotic attributes and empathetically co-enacted perceptual attributes of VDM per se are projected onto the overall aesthetics of the multi-modal phenomenon. The narrative is thereby shaped through its affiliation with these attributes. In other words, when speaking of computer games it can be said that the general stylistic expression of VDM influences the general stylistic expression of the game as a whole. This general relation of VDM to a game partly accounts for how VDM may narrate in such a context.

Additionally, *specific* music-structural constituents of VDM such as certain instruments, playing techniques, registers, timbres, textures, harmonies and other vertical parameters (see list of vertical parameters in Appendix F) or horizontal development schemes (horizontal development in VDM is covered in a chapter 9) may, according to how they are implemented in, for instance, a computer game, be coupled to *specific* in-game elements such as characters, locations or situations. Here the musical sub-structures constituting the musical expression, which can be set as manipulable parameters in a generative system, each have an individual narrative potential. This way for VDM to narrate is different from the general relationship mentioned above by being more closely tied to the immediate action of the game. It is on one hand concerned with projecting the general semiotic and perceptual attributes of VDM onto the specific situation at hand in the game, while on the other hand it assigns concrete sub-structures of the vertically dominated musical expression (such as erratic string tremolos in a deep register, a cluster of non-vibrato flute tones at a high register or a synthetic harmony-timbre modeled on the overtone distribution of whale song - all of which may be sounding simultaneously) to individual components of the game narrative. It seems evident that if this is to work (i.e. if individual elements of a vertically dominated musical expression is to be used narratively by coupling each to their own in-game narrative component) then the vertical expressions of the music should allow for what Albert Bregman calls *sequential* and *simultaneous organization* (discussed in chapter 2). In other words, based on the example above, it should be possible to distinguish the strings from the flutes and the synthesizer.

Furthermore, the specific semiotic and perceptual attributes of such vertically dominated musical sub-structures may be already established as a culturally conditioned media competence (such as dissonant tremolo strings being stereotypically associated to some sort of danger, while a mellow string chord might be associated to feelings of love and so on). One might refer to such competences as *game-external* as they are broadly culturally derived rather than initiated by a particular game being played. Additionally, meaning can be induced in specific musical sub-structures by way of game specific *priming*[36] (Tulving & Schacter, 1990) of the player - creating a *game-internal* decoding competence where players learn the meaning of the music through playing the game (such as if a noisy synth pad is repeatedly coupled to the same game event). Such priming fuses the newly constructed game-internal sign with any sign held in the game-external competence of the player. The conditioning that informs the competence of the player can thus be based on either schematic or veridical memory or both. In effect, some usages of sub-structures as game-internal signs will correspond better than others with game-external competences, increasing the likelihood of the sign being decoded appropriately.

6.3.4. CONCEPT DEVELOPMENT FOR TRANS-DIEGETIC CROSS-MODAL CONNECTIONS

The specific coupling of musical sub-structures to in-game components is a question of design. Such a design requires the development of a concept in which game-external and game-internal competences, which are drawn upon in the creation of musically facilitated meaning, are in line with the overall game design strategy.

Building on the vertical parameters, which were categorized in chapter 3 into parameters of tone system, harmony, timbre and texture, a cross-modal coupling can be organized according to narrative and aesthetic aims between the adaptive VDM and other elements of the gameworld, game arena, gameplay and game narrative. In a FPS game (First Person Shooter) the register and instrumentation of the generated music might for example be mapped to the number of enemies within a certain distance from the avatar. In an MMORPG (Massively Multi Player Online Role Playing Game) like EVE Online or World of Warcraft the amount of spectral compression or stretching might be associated with the hierarchy of power of different in-game NPCs (Non-Player-Characters) and so on.

Somewhat similar conceptualization considerations are related to the film medium. While it can be problematic to project theory and work methods from the film media directly onto the game media due the latter’s non-linearity, in the case of conceptualization much can be learned from film. This is not necessarily problematic because the music’s underlying principles or concepts (i.e. the cross-modal coupling strategies between music and other elements of the media) lie at a level prior to implementation. In games, however, such a conceptualization task has to be extended to take into account also the music’s adaptive implementation, functionality and trans-diegetic potentials as well as bear in mind a number of

---

[36] *Priming* refers to ”…a change in the ability to identify or produce an item as a consequence of a specific prior encounter with it.” (Calvert, Spence, & Stein, 2004; Tulving & Schacter, 1990)

technological restraints. Such issues are discussed in chapter 7, which also looks at possible narrative functions of game music.

In film music, the conceptual design behind the film score lies at the core of the film-musical expression. In western filmic tradition, the underlying concept represents the primary idea, the foundation on which the whole score is based, and musical consistency towards the developed concept in a film is important in building and maintaining dramatic integrity (Karlin & Wright, 2013, pp. 66-110). The musical concept plays an important role in outlining the design strategy in regard to thematic material, musical style, instrumentation and orchestral color and can be focused towards narrative elements such as its central characters, overall dramatic scene (or scenes), which sums up the movie, as well as ethnic and geographic attributes of the narrative to the desired extent. When focusing on one or more central characters, the strategy can be aimed at for example the character's background, situation in the drama, emotional state, personality or state of mind. The character can sometimes be unseen in the movie in which case the music takes on the important role of implying its presence. The character might also be inanimate as in *"Close Encounters of a Third Kind"* (Spielberg, 1977), where the music expresses qualities of an extra-terrestrial space ship as it is experienced by two central human characters (Karlin & Wright, 2013, pp. 66-110).

Film music based on a concept built around the overall dramatic scene of a film tends to stand back from the immediate action and instead take on a role of reflection and overview. Ethnicity and geography can be suggested through instrumentation, harmony, motifs, rhythms and scales typical of a certain area or culture. Many times such a focus is based more on stereotypical western prejudice than on what music is actually predominant in whatever culture is being suggested, since this music in reality may very often be not too different from the music of the western world due to globalization. And in many cases ethnic and geographic elements in a film score function mostly on an evocative level in a score with otherwise traditionally western musical stylistic traits (Karlin & Wright, 2013, pp. 66-110).

The conceptual choice of musical style lies beyond the focus of the present thesis as the musical style under the loop here is predefined as VDM. However, within VDM certain styles can be identified as was mentioned in chapter 4 (e.g. sound-mass, spectral, ambient and noise music). The focus of my research is aimed at making such conceptual choices possible to implement adaptively within an interactive computer game context. These choices may include symphonic or electronic instrumentation, why the capabilities for manipulating VDM's structural parameters should encompass at least both of these instrumentation paradigms by offering both sample and synthesis based operation. I will go into these questions more elaborately in chapter 8.

HDM's linguistic musicality (i.e. the hierarchical organization of low-level units into higher level units and the causal relations between them) associated with causal direction on the levels of harmony, melody and rhythm with its potentials for narration through these phenomena, is ideally absent in VDM. So, too, is narrative possibilities offered by the tensions of the harmonic system of tonal music, which is so well established in film and game music. A concept within a vertically dominated musical stylistic frame will thus have to replace such horizontally dominated musical parameters as a basis for conceptual musical effects with the vertical structural parameters identified in chapter 3 and 4 and listed in Appendix F. Because VDM in its purest form is devoid of repetition, traditional narrative concepts such as for example the *leitmotif,* which is so effective and widely used as a cross-modal narrative technique in opera, films and games, may be replaced by vertically dominated and horizontally less specific structuring schemes. These could look something like [tremolo strings, forte, low register], [square wave synthesizer, intense pitch modulation with depth of 1 semitone, high-pass filter cut of at 2kHz, 2nd harmony, medium register] or [brass section, piano, wide hollow cluster] and so on.

The design of such a conceptual mapping of vertically dominated musical structure to specific game elements may draw inspiration from well established film and game musical stereotypes as well as making use of game-specific priming - thus taking into consideration both game-external and game-internal signification.

6.4. OUTER SPACE AS NARRATIVE SETTING

Having established that VDM may narrate on the basis of competence-related semiotic attributes tied to the style's prior association with scenarios of horror and mystery as well as narrate through empathetically co-enacted perceptual attributes, which I have condensed to the adjectives *continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless* and *unpredictable*, it is now time to take a look at VDM's connection to the narrative setting of outer space. This enquiry highlights indications that suggest a correlation between also the *perceptual* attributes of the sound and music that has been traditionally associated with outer space in science fiction and the perceptual attributes of VDM.

As I mentioned in the introductory chapter, the coupling of outer space and VDM in films like *"2001: A Space Odyssey"* is what first triggered my interest in the subject matter of the present research. Also, my ongoing collaboration with CCP in regard to their space-based MMO, EVE Online, and the use of this game as a case for practical experiments presented in Appendices A, B, C, D and E, serve as a natural motivation for looking into this narrative setting in some detail.

The *existential mystery*, which can be associated with outer space, may be seen, however, as a common denominator to also the other narrative phenomena included in the equities of this thesis (infinity, mystery, fear, divinity, paranormal activity and a presence of "something" that exceeds the boundaries of the senses) and findings from this focus are easily transferrable to these other phenomena. I thus use outer space here as an example that will allow for a deeper look at how VDM may narrate on a more concrete level. The focus on outer space is thus mostly a matter of focus and is not an expression of this narrative phenomena being in any way closer related to VDM than the others - neither from the perspectives of semiotics nor embodied experience.

As only very few people have actually been outside the Earth's atmosphere, most people's conception of outer space builds on a mixture of popular science, looking at the stars at night (or rather the space in-between the stars), and ideas more or less influenced by fiction in one form or the other. A significant part of the science fiction genre in literature, film and games takes place in outer space. Pictures from distant galaxies taken by the Hubble Space Telescope is imprinted in the memory of most of us, just as it can no longer be considered uncommon to see pictures of our own Earth in its entirety from outer space. This was obviously an immense sensation around the cradle of space travel in the 1960's. To see Earth ”hang” there alone in the emptiness of the vast and dark vacuum for the first time must have made an indelible impression on those who were lucky to be there for this premiere.

However, a certain trivialization of outer space has arguably taken place through the huge media exposure it has received. This trivialization is reflected in plot and focus of some of the ”space films” that have been released. In many cases - such as in the *"Star Wars"* films (e.g. Abrams, 2015; Lucas, 1977), *”Armageddon"* (Bay, 1998) and *"Aliens"* (Cameron, 1986) to name a few - there is a greater emphasis on heroic tales of action and adventure than on the philosophical and existential implications of the extraterrestrial scene in which they take place. Films such as *"2001: A Space Odyssey"* (Kubrick, 1968) and *"Solaris"* (Tarkovsky, 1972) from the late 1960's are examples with a stronger focus on these aspects, but they are not alone on the existential, philosophical or mystical focus as evidenced by newer films such as *"Event Horizon"* (Anderson, 1997), *"Contact"* (Zemeckis, 1997), *"Sunshine"* (Boyle, 2007), *”Gravity”* (Alfonso Cuarón, 2013) and *”Interstellar”* (Nolan, 2014) where mystery and existential implications play a significant role in the narrative.

While an investigation of accompanying sound and music in a large number of space scenarios in film and games would be useful, presenting a comprehensive study of what the media competence of the film audience and gaming community might contain is outside the scope of the present research's prioritization. Such a study might give a precise account for (or at least a qualified hint of) what gamers and film audience from varying cultures and segments are used to in regard to sound and music accompanying deep space in space fiction. A study of music in *fantastic* contexts in film, including space scenarios, can be found in (Larson, 1985) as well as (Bartkowiak, 2010).

Although I will not here provide a historical outline, a few interesting aspects from the history of space science fiction film music are worth mentioning - especially the relationship between the narrative scenarios of *outer* and *inner space* and their relation to the genre's music and sound design. This is relevant not least because VDM's association with outer space is partly responsible for the *game-external* competence of this music in terms of *general* narrative couplings.

That there is a historic correlation between VDM and space based science fiction, horror and mystery genres is evidenced by genre classics such as *"2001: A Space Odyssey"* (Kubrick, 1968), "*Solaris"* (Tarkovsky, 1972), "*Alien"* (Scott, 1979), *"The Shining"* (Kubrick 1980) to name a few of the earlier films that were responsible for solidifying the style's symbolism. VDM figures to some extent, in almost every horror and science fiction movie made since, in some cases only as an effect accompanying particularly scary or shocking passages, other times serving as the musical backbone of the film's non-diegetic narrative layer.

6.4.1. SOUNDS OF THE GOLDEN AGE OF SPACE-ORIENTED SCIENCE FICTION

In space movies from the 1950's, however, the music took on a different approach to conveying this inhospitable narrative setting. While this approach may be different, it bears important similarities with the VDM now so widely used. James Wierzbicki in his article "The Imagined Sounds of Outer Space" (Wierzbicki, 2014) examines the relationship between a number of narrative elements of outer space on film, such as space travel and foreign planets, and the sounds accompanying these elements. The following section highlights, very briefly, some of his findings with the aim of giving an insight into the historic development of the media competence of space oriented film audiences.

In American, Soviet and European space films from the so-called Golden Age of space-oriented science fiction in the 1950's, the sounds accompanying planets, space travel and space signals were produced by the most recent technology of the time, and often these futuristic sounds were based on models, well known to the public. The sound design for signals in outer space could, for instance, be based on communication sounds as they took shape on earth, imitating morse code and short wave radio. Sounds from space travel were modeled mainly on sounds one could expect to hear on an airplane. The result could sometimes be of a somewhat mundane character.

The soundtracks for these films, however, also featured non-traditional musical instrument technologies such as the Theremin and Ondes Martenot as well as the perhaps lesser known Trautonium and the Russian ANS synthesizer. But the use of, for instance, the Theremin as an artifact of film sound and music originally began in the 1940's, coupled not to science fiction but to abnormal states of consciousness - to start with in association with scenes of psychoanalysis in *"Lady in the Dark"* (Leisen, 1944) and Hitchcock's *"Spellbound"* (Hitchcock, 1945). Almost at the same time in *"The Lost Weekend"* (Wilder, 1945) theremin was used in scenes following a raving alcoholic. In a period between 1944-1951 the theremin figured heavily as a film sound effect to convey everything from paranoia, hypnosis, nightmares and drug use - effectively establishing it by the early 50's as a symbol for an unbalanced inner psychological state in the media competence of the film audience of the time. It wasn't until *"Rocketship X-M"* (Neumann, 1950) that the theremin saw its first appearance in a space based science fiction scenario. In the following years, the theremin and many other electronic instruments went through a shift in audiovisual semantic meaning as they were increasingly used in the context of space oriented science fiction.

It is interesting to note that the use of this equipment meant for music and the implementation of said instruments in connection with especially space travel, often resulted in a somewhat blurry border between the diegetic and non-diegetic narrative layers because you couldn't be completely sure as an audience, whether or not what you were hearing was meant as accompanying music or coming from an extra-terrestrial source inside the reality of the narrative. This is the case in films such as *"Forbidden Planet"* (Wilcox, 1956) and *"Ikárie XB-1"* (Polák, 1963). (Wierzbicki, 2014).

The sound design and music strategy of the 50's has at least a few things in common with the principles of VDM. They are linked by their uncontrollability, unstable pitch and by not conforming to a traditional musical logic. They are capable of presenting a disintegration of order by violating pitch boundaries associated with acoustic instruments and by having potentially infinite tone duration. Furthermore, they can be seen to convey an *otherness* by being uncommon to the ear of the time.

In regard to music, this otherness, strange and alien quality, has long been tradition in space-oriented science fiction, and it is fair to assert that these attributes are present in the media competence of the general film audiences and gamers.

6.4.2. INFINITE REVERB AND SOUND IN SPACE

Obviously, when dealing with outer space there are no walls on which sounds can ricochet, nor any medium for audible sound to travel in. The problem of diegetic sound in space scenes has been approached in a variety of ways spanning from: A complete disregard of the impossibility of sound in space (e.g. the *"Star Trek"* series and, in fact, "*EVE Online"*); to a focus on the subjective auditive experience of characters inside their space helmets or within the hull of a spacecraft (e.g. *"Sunshine"* (2007), *"Gravity"* (2013) and *"2001: A Space Odyssey"* (1968); to occasionally maintaining complete silence (e.g. *"Gravity"* (2013) and *"Interstellar"* (2014).

The problem of sound in space can seem unimportant in regard to non-diegetic sound and music as these do not have to subjugate to the logic of the diegesis. But when outer space takes the role of the *undisclosed sender* (not a sound source) the problem arises of how a phenomenon such as space, which is essentially "nothing" in itself, can be associated with a *state* to be conveyed. Space is a dimension, not an object that may easily be attributed a set of qualities. However, perceptual principles that are violated with the introduction of a sort of unlimited spaciousness achieved aurally through a seemingly unlimited reverberation tail made possible by audio technology, may allow the non-diegetic narrative layer to convey information of the *state* of outer space on a meta-level. Keeping in mind Deleuze and Guattari's elaboration of Boulez' notion of smooth and striated space-time discussed in chapter 2, ”infinite reverb” may be understood as an expression of ”space” rather than ”room” in the sense that a room is essentially a striation of space, while space itself is the dimension in which room occupies - space is unstriated and intact in its infinite continuity. As such, this special form of reverberation plays with some of the same perceptual attributes as those of VDM, that I have discussed earlier (such as *continuous, limitless, timeless, infinite,* and *incomprehensible).*

We are familiar with hearing very long reverberation tails from large acoustic cavities such as cathedrals, caves and huge warehouses, but at some point the duration of the reverberation's decay becomes super-natural by exceeding the boundaries of our experience with realistic space. Infinite reverb thus facilitates on a spatial basis some of the same perceptual attributes associated with VDM also in regard to a breaking with logic.

Traditionally, ”infinite” or extremely long reverb tails have been used in science fiction films to convey a sense of loneliness and emptiness (e.g. the opening scene of *”Alien”* (1979)). This effect could, for instance, be achieved by filtered noise or other eerie sound sources accompanied by images of the vast emptiness of outer space, empty hallways, barren deserts and so forth. This audiovisual expression is typically part of a narrative strategy to provide a sense of loneliness or desertion - drawing upon, perhaps, an inherent human fear of isolation. The sense of size of this imaginary space is induced by a sense of distance to the sound source caused by the reverberation and creates depth to the sound stage. In turn, the sonorous characteristics of the sound source may be obscured to the extend that source identification becomes impossible. In fact, in many cases the role of the sound source is exactly to be diffuse and undefined while, more importantly, providing a means of articulating the imaginary non-diegetic resonating space. In these instances the ”space” in itself, rather than its content, is key, letting anonymously articulated infinite-scale reverberation exceed our capacity for spatial comprehension.

6.5. BENIGN VDM

A special usage of consonant VDM is emerging in films like *”Maggie”* (Hobson, 2015) and ”*Synchronicity”* (Gentry 2015). Here images of events, seemingly unimportant to the storyline (such as flowers moving slowly in between the sun and the camera lens producing a flare, images of snow slowly falling, dust floating in the air etc.) are presented often in slow motion accompanied by consonant but clearly vertically dominated music. This cinematic style has a calming effect and seems associated with stepping out of the narrative for a moment of reflexive hindsight or in order to let the emotional situation of the movie sink in. While perhaps most often encountered in commercials, some newer feature films use this type of audiovisual expression quite a lot, and one can occasionally wonder whether it is sometimes used to fill in the gaps in a narrative that is coming short on content for a full feature length movie.

Significantly, this type of VDM usage can be seen as yet an example of the style's coupling to something *beyond* the diegesis. By turning the focus towards all that which no-one (in the diegesis) is experiencing, or which is experienced by a character in a meditative state, the emphasis is momentarily placed on a sort of timeless dimension - on events, which have no immediate importance and entail no causality in terms of the narrative's further development. The intention of the film makers in such cases seems to be to promote, in the audience, a meditative state of consciousness through an emphasis on a timeless ”now” that is beyond the drama itself. It represents a positively valenced - or *benign* - form of VDM, which may hold an equal narrative potential as the more dissonant and sinister VDM.

As discussed in the previous chapter in the section ”Movement at the event horizon”, this implies that the *continuous, limitless, uncontrollable, chaotic and timeless* qualities brought about by *vertical dominance* synthesize with the semiotic attributes of the *vertical expression* in whatever form the vertical expression may have. Consonant and friendly sounding vertical expressions thus become friendly with the added qualities of timelessness, continuity, infinity and so on, with a potential for promoting associations to phenomena like divinity, bliss, deep relaxation and meditative states of consciousness.

6.6. INNER/OUTER SPACE

An endeavor to convey through music the idea of outer space, which bears in mind the notion of media competence, must take into account the audience's and player's existing connotations of outer space. These connotations come to exist in the receiver's consciousness. As such, the music must describe the extrinsic outer space as it is represented in an intrinsic inner space. Setting aside the influence of cultural myth and religion, from the perspective that sees musical listening as a process of empathetic co-agency, one could thus argue that the decoding that creates the connotations of *outer* space is based on the conditions of an *inner* mental space. That is, cognitive schemas for motor-control and physical limitations of the human body, the evolutionary physical and psychological conditioning to the laws of nature as they are on Earth including a gravitational pull of one *G* and the boundaries in (or striation of) the spatial dimension by the ground, objects and life forms, the presence of air, pressure of one atmosphere etc.

For most humans, who are conditioned by this cognitive schematic framework of comprehension, the phenomena of *outer space* and the *Universe* are impossible to grasp. However, outer space attributes like *infinity* and *incomprehensibility* in themselves create the very foundation of man's limited realization of this great mystery. Thus outer space is paradoxically perhaps best conveyed musically by bringing into play such preexisting embodied experiential structures of inner space - by juxtaposing the condition of our embodied schemas with musical structures designed to break the boundaries of such schemas through empathetic co-agency.

Our individual cognitive schematic framework for music - our musical *history,* as Ligeti puts it (see chapter 5), which facilitates specific musical expectations serves a narrative function of acting as a model, a symbol, for other elements of a multi-modal medium. But, and this is an essential point, embodied experience implies that breaking with this culturally based and top-down processed framework also entails an ability to break with the cognitively fundamental and natural bottom-up processed preconditions of human perception.

The *continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown,* weightless*, homeless* and *unpredictable* character of VDM and this music's ebb and flow of *unrestricted, unstriated, continuous* movement corresponds well with semiotic attributes of outer space. Precisely this musical connection to the movements of *inner* space as a base for humans' realization of *outer* space is arguably what allows VDM to work so well for this task.

These characteristics of VDM make it well suited, not only in space oriented science fiction, but also in horror and mystery genres, as well as, in its benign form, in scenarios of meditative peace - all of which may be regarded as examples of a presence of ”something” that exceeds the boundaries of the senses.

6.7. CONCLUSION

In this chapter, I have investigated how and what VDM narrates. I have enquired into this from four different perspectives:

- The role of different diegetic layers in the formation of multi-modal meaning
- Competence-related semiotic attributes of VDM
- Empathetically enacted perceptual attributes of VDM
- VDM's relation to its multi-modal context

Furthermore, I have related these four aspects of VDM narration to outer space as a narrative setting and through this drawn parallels to other phenomena of interest in the thesis, such as infinity, mystery, divinity, paranormal activity and a presence of ”something” that exceeds the boundaries of the senses.

The concept of diegesis and especially the distinction between the diegetic and non-diegetic is problematic in regard to computer games. I have argued, however, that from a phenomenological perspective, the media competence of players may be seen to dictate its presence in film-like games as it has been projected from the film media. The inherent mystery of the non-diegetic layer, which I referred to as ”the great beyond" of the non-diegetic, can be said to further support the "undisclosedness" of non-diegetic music's undisclosed sender in cases where a linear sender-music-receiver communicational paradigm is phenomenologically maintained because any trans-diegetic functionality is *covert*. This effect of supporting the strength, so to speak, of the narrative potential of non-diegetic music, may on the other hand be weakened in cases where a non-linear, circular communicational paradigm exists due to a phenomenologically *overt* trans-diegetic functionality.

I have argued that VDM, and music in general, may be seen to narrate through what I call *competence-related semiotic attributes* and *empathetically co-enacted perceptual attributes* that are projected onto whichever element of the multi-modal construct the music couples to. For VDM this means that the signification which the style has received in the competence of audiences and players through its association with horror and mystery scenarios, will likely cause such associations to underscore any focus on horror and mystery in new games featuring VDM. Furthermore, the perceptual attributes, which I condensed to the adjectives *continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless* and *unpredictable* earlier in the thesis, are arguably likewise projected onto other elements of a game. Thus, in addition to the semiotic attributes, VDM may also narrate by virtue of the cross-modal coupling of this set of perceptual attributes to other components of the multi-modal phenomenon.

Such narrative couplings may happen in a general or a specific manner. A *general* coupling implies that the mentioned semiotic and perceptual attributes of VDM are projected onto the *overall* narrative and aesthetics of any multi-modal phenomenon that features VDM. As another way of narrating musically with VDM, *specific* music-structural constituents of the style (such as certain instruments, playing techniques, registers, timbres, textures, harmonies and other vertical parameters) may be coupled to *specific* in-game elements such as characters, locations or situations. This *specific* approach to VDM narration in multi-modal contexts requires a degree of vertical separability and the application of the auditory gestalt principles highlighted by Bregman and McAdams and their theory of auditory stream analysis may provide valuable insights that can be used compositionally to achieve this goal. Furthermore, specific semiotic and perceptual attributes of vertically dominated musical sub-structures may be subject to veridical *game-internal* and schematic *game-external* competence, drawing on game specific *priming* or broadly culturally derived expectations respectively. Some usages of sub-structures as game-internal signs will correspond better than others with game-external competences due to a good correlation between the two. This may increase the likelihood of such signs being decoded appropriately according to the intentions of the game design.

The design of such specific couplings are ultimately conceptual choices. I have highlighted that the development of a VDM based game-musical concept has to build on different musical structural schemes than HDM, as traditional techniques such as leitmotif and repetition in general are absent in the style. These other structural schemes, which are constituted by the list of vertical parameters highlighted in chapter 3 and 4, include instrumentation, playing technique, amount of dissonance, degree of spectral stretching and may be devised by any of the vertical parameters of Appendix F. Such vertically oriented structuring schemes may serve as vertically dominated equivalents to the horizontally dominated notion of the leitmotif, or they may serve to perform other game-musical functions (a list of game-musical functions are taken up in the next chapter). Solutions for the development of vertically dominated concept designs in this sense may be anchored game-internally, game-externally or, preferably, in both.

Sound in space-oriented science fiction has long relied on ”strange” sounds to convey other worlds. VDM holds in common with, for example, the use of theremin in the 40’s and 50’s, a tendency to have been used to convey both altered states of consciousness as well as outer space as a narrative setting and the presence of ”something” that exceeds the boundaries of the senses - or, which in other words is part of the diegetic world, while not being *of* it.

Furthermore, VDM’s association with scenarios of horror is especially characterized by dissonant and negatively valenced vertical expressions. Additionally, this narrative potential of VDM is countered in consonant forms of *benign* VDM to convey meditative peace and reflection.

Seemingly infinite and super-natural reverberation tails may allow the non-diegetic narrative layer to convey information of the *state* of outer space on a meta-level even though *space,* being a dimension, doesn’t have any inherent qualities of its own that are easily conveyed or understood. Infinite reverb can be seen as an expression of unstriated ”space” rather than striated ”room” and in this way play with some of the same perceptual attributes as those of VDM such as infinity, incomprehensibility, continuity and limitlessness through a breaking with logic and corporal limits of embodied experience. This form of reverb holds the potential for letting ”space” itself, rather than its content, be key - freeing it to exceed our capacity for spatial comprehension.

VDM, as experienced through empathetic co-agency, can be argued to break with the boundaries of an *inner space* in order to convey *outer space.* This represents a special narrative ability of VDM that is of great significance in the context of this thesis. Importantly, I have thus been able to identify that breaking with the culturally based and top-down processed framework of our musical competence through vertical dominance, also entails an ability to break with the cognitively fundamental and natural bottom-up processed preconditions of human perception. And it is not least this special ability, I argue, that lets VDM so effectively give a voice to such phenomena as outer space or a presence of ”something” that exceeds the boundaries of the senses.

As has been discussed in this chapter there are significant similarities between the film media and computer games in the way they narrate - especially on a conceptual level - and a complete separation of the two as individual autonomous media is too radical, unnecessary and ultimately disregards the possibility of a projected media competence. The subject matter of the next chapter is the challenges that are nonetheless posed by the computer game media by its unique characteristics. These will include issues concerning non-linearity, immersion, presence, various genres such as MMOs and FPSs and the different challenges that reside with these as well as technical and functionality issues.

CHAPTER 7. SOUND AND MUSIC IN THE COMPUTER GAME MEDIA

VDM's abilities for narration, as discussed in the previous chapter, are primarily facilitated by semiotic and perceptual attributes. Some of these attributes are common to also the film medium. Although some games appear much like films you can play along in, many games do not, and a range of issues set the computer game medium apart from the film medium and other linear media. It is the purpose of this chapter to cover some of the most important of such aspects relating to music and sound.

7.1. INTRODUCTION

So far this thesis has centered mainly on VDM, its philosophical, structural and narrative aspects. It is now time to investigate another crucial pillar in the foundation for a generative compositional system of VDM to rest upon, namely the implications related to the medium in which it must operate. Any initial set of compositional principles of such a system must take into account basic aspects of the computer game medium as it represents both challenges and opportunities that cannot be ignored. The framework for understanding VDM's impact on the gaming experience is thus further elaborated in this chapter as it connects VDM structure and perception to the general challenges of game music and sound - determining the game media's potential impact on the music and vice versa.

The majority of the present chapter deals with the identification and discussion of a number of issues and challenges confronting the game music composers and sound designers. These issues and challenges are in turn related to the unique characteristics of VDM in order to present an overview of how VDM as a music compositional technique with its unique structural and perceptual qualities may be expected to perform in the interactive environment of the computer game medium.

The chapter is structured into five main topics titled: General challenges of computer game sound; Adaptive music; Immersion and avoiding competition for player attention; Functions of game music; and VDM and the challenges of computer game music.

Firstly, I shall highlight some general challenges that relate to both sound and music in computer games. These include implications in regard to the technological frame, game genres and not least the non-linearity of the medium. Technological advances have expanded the possibilities for music composition in games, although the limitations posed by hardware are still today a key aspect of sound in games. I present here an overview of significant technological implications. The computer game medium is very diverse and different game genres pose different challenges to game design, sound design and music composition. This chapter presents a look at some genre specific challenges with special emphasis on MMO, sandbox and persistent universe style games. The non-linearity of the game medium refers to an indeterminacy in regard to the course of events in the games' narrative. The notion of non-linearity in games can be criticized, however, for being too categorical for a medium that, in most genres, exhibit some form of linearity. The term is nevertheless useful to denote the obvious difference to clearly linear media caused by its inherent interactivity.

Such non-linearity calls for music and sound which are likewise non-linear in structure in order to synchronize with the action of the game. This is discussed in the second topic titled *adaptive music*. Here I discuss the notions of *dynamic*, *interactive* and *adaptive music* and propose the terms *direct* and *indirect adaptive music* inspired by Karen Collins' notions of *interactive* and *adaptive dynamic music*. From here I go on to describe some common techniques associated with the implementation of adaptive music such as branching, layering, transition matrixes and generative music. I pay special attention to how these techniques may differ in game designs and musical styles that have either an arborescent or a rhizomatic structure. The discussion of these concepts leads me to arguing that VDM, being structurally comparable to rhizomes rather than trees, fits game designs with similarly rhizomatic structures - such as ”persistent universe sandbox MMOs” – very well in terms of implementation. As a last part of the second topic I look at challenges and solutions for creating variation in game music and present a VDM specific list of approaches to musical variation in games.

As the third topic, the notions of *immersion* and *presence* are briefly discussed due to their significance to the quality of the gaming experience. While generally pursued by the game industry as a criterion of success, I argue that the significance of immersion and presence in games varies according to game type. I further propose that in an effort to avoid competition for player attention, and thus a weakening of immersion, music that is heard at a listening attention level, referred to by Barry Truax as *background listening* may pose less of a threat to breaking a player's immersive state than the listening attention levels of *listening-in-search* and *listening-in-readiness.*

Adaptive music in games may take a range of different functions. I discuss possible game-musical functions inspired by not only music but also film- and game sound design. This section of the chapter goes on to arguing that music in games may perform any film-musical function, take on functions normally attributed to sound design as well as perform trans-diegetic functions specific to the computer game medium. A list of game-musical functions is presented.

Finally, the fifth topic relates the findings of the chapter to the structural and music-perceptual properties of VDM and presents an overview of the implications of implementing a vertically dominated approach to music composition in computer games.

The chapter concludes by arguing that the unique structural and perceptual attributes of VDM places it in an advantageous position as compared to HDM in regard to all of the above mentioned game related challenges - except for the existence of some technological challenges, not least associated with VDMs tendencies for very large-scale polyphony and demands for a high degree complexity in its *vertical expression*.

7.2. GENERAL CHALLENGES OF COMPUTER GAME SOUND

7.2.1. TECHNOLOGICAL FRAME

Since the birth of computer game sound in the arcade games of the 1970s, game programmers and composers have had to deal with the challenge of taking full advantage of the technological possibilities offered by the hardware of their day. A constant balance have had to be found between the creative agenda and the limitations put by technology to unfold this creativity. Although the boundaries of what is possible technologically have moved a lot since the 1970s causing a larger creative leeway, the technological boundaries are set today by practically the same barriers as then. The sound chip's or audio interface's capacity for sound reproduction, the amount of disc space and the amount of memory and processing power available for game sound still stand as bottlenecks through which the creative output of the composer and sound designers must be able to pass. Already in the 8-bit era[37] composers, who at the time were often programmers, thus developed certain techniques designed to get as much music as possible out of the technological frame at hand. These techniques include the implementation of musical material through loops, random sequencing, and algorithmic composition (K. Collins, 2008, ch. 2), of which the latter allowed the system to compose music based on a programmed algorithm and played through an onboard synthesizer.

The same techniques are principally in use today. The difference between the present and the 8-bit era lies more in an increase in the efficiency of the techniques with a resulting refinement of expressivity rather than in the techniques having been substituted for fundamentally new approaches. Thus, certain side effects of dynamic music have been met, as is the case with the development of a higher diversity of transitions between musical sequences, an increase in the number of available loops and possibilities for creating more complex and thereby more technologically

[37] Refers to a period in the mid 1980s dominated by 8-bit video game consoles. Described in (K. Collins, 2008).

demanding compositional algorithms (which now often manifest sounding music via sample libraries), the availability of more musical layers and timbral variants.

7.2.2. A SHORT NOTE ON COMPUTER GAME GENRES

Some games have a particularly high degree of non-linearity. To distinguish these games from others, I therefore include here a short section on computer game genres. As is generally the case with qualitative and historical categorizations also in regard to game genres the categories are subject to somewhat floating boundaries and large grey zone areas in which we in the name of clarity place divisions in order to create an overview and be able to say something about a game's character with very few words. Nonetheless, such divisions can be useful in the placement of a particular game in its broader context of the computer game medium - although such a placement must be considered somewhat diffuse and subject to change over time. As with genre categorizations in for instance modern electronic music, for games the labeling is governed by a high degree of autonomy as well as a limited but to a certain extent existing consensus on the categories.

The Entertainment Software Association (ESA) puts forth a number of so-called ”super-genres” in their statistics. These include action, adventure, arcade, casual, family entertainment, flight, racing, role-playing, strategy, sport games and shooter (ESA, 2015), but many other ways of categorizing exist. Often a game's affiliation to one category or the other is assessed differently on individual computer game-related websites. I will exemplify this with three prominent game-related sites: mobygames.com, ign.com and mmorpg.com - although several others might equally well deserve mentioning. Let us take as an example EVE Online, the Icelandic game that I have used as a case in my research for this thesis. EVE Online is categorized by mobygames.com under the umbrella of RPGs (Role-Playing Game). The site additionally positions the game under the labels 3rd Person Shooter, Managerial, Persistent Universe and Science fiction / Futuristic (mobygames, 2016). On ign.com, EVE Online is archived under Persistent Online RPG (ign, 2016). At other places the game is referred to as a sandbox MMORPG (Massively Multiplayer Online Role-Playing Game) (mmorpg, 2016). As illustrated by these genre names the aspects of emphasis in the categorizations differ. FPS denotes the visual perspective from which the player sees the game world. Persistent Universe Online RPG means that the game is a role-playing game taking place in a game world, which is running constantly - also when the player is not actively playing the game or even has the game program opened on his or her computer or console. The MMO category refers to the fact that a game is played by many players over the internet at the same time and that the players play together in the same[38] game world. Other game sites use yet

[38] According to CCP EVE Online is presently the only large scale MMO to offer one consistent universe for all players. Games such as World of Warcraft have players divided into smaller game worlds of a few thousand players in each.

other genre labels, but within the frame of this thesis, it is sufficient to note that EVE Online can be categorized as a sandbox MMORPG with a persistent universe.

Such games have particular properties in regard to the degree of non-linearity and the compositional challenges this entails as will be more evident in the following pages.

7.2.3. NON-LINEARITY

A fundamental aspect of the implementation of music in games is concerned with the non-linearity of the computer game medium. Non-linearity is here understood as an absence of a predetermined (and in that sense *linear*) narrative progression that governs for example the film medium and allows precise synchronization and coherence between the narrative expressions of sound and image - an absence, which poses a fundamental challenge to the game music composer. The fact that a great deal of control over the course of events is in the hands of the players means that linear music and sound is most often only to a limited degree capable of functioning as anything more than a background tapestry for events occurring on the screen. Linear music accompanying a non-linear course of events is largely decoupled from the game's virtual reality because any coherence between music and game events are left to chance. (Unless the player chooses to synchronized his or her actions to the music, which, as I have mentioned earlier, is the very point of music games like *Rockband* and *Guitar Hero*.)

However, computer games are not necessarily non-linear all the time. In some games the gameplay (or part of it) is based on a defined period of time - like when the player is presented with the task of reaching the next checkpoint for instance in a racing game, or has exactly 3 minutes to destroy an open office environment before the guards arrive. The fixed time period of a ”round” in a boxing game is another example. So-called *cinematics* or *cut-scenes* also introduce linearity in some games. Here the interactivity of the game is interrupted for a brief time by linear animated sequences that are often used to move the action from A to B - for example between game levels. While the game medium, due to its interactive nature, may appropriately be considered non-linear, it should be noted that this non-linearity is scalable - that the degree to which games are non-linear varies. Typically, there is an aspect of predetermination in the progression of a game. It may be player-controlled exactly *how* the action moves from scenario A to scenario B, but that it *will* move from A to B (as well as what these scenarios entail) is in most cases a matter of game design and arguably an expression of linearity. There is for instance a greater degree of non-linearity in the course of events of a persistent universe sandbox MMO like EVE Online than that of a puzzle game like Portal (Valve, 2007), where the player has to perform a series of tasks in the right order to move to the next (predetermined) game level. One may therefore criticize the term for being too absolute when generally applied to games.

Nonetheless, non-linearity and indeterminacy per definition play a role in interactive media giving sound designers and composers only to limited extents a predefined reference for the synchronization of image and sound. Such media demand that music and sound be also presented in a non-linear way.

7.3. ADAPTIVE MUSIC

A large body of academic work exists on game sound, while the subject of game music is somewhat more scarcely covered. In this section I examine a number of common aspects of game music and place them in the context of VDM.

7.3.1. DIRECT AND INDIRECT ADAPTIVE MUSIC

In order to address the indeterminacy associated with a non-linear course of events, very often computer games will feature an implementation of sound and music that is capable of adapting to the in-game scenarios as they unfold. Karen Collins calls such sound and music *dynamic*. Collins writes about dynamic music and sound that it can be functionally connected to the gameplay either *interactively* or *adaptively* (K. Collins, 2008, pp. 4). Interactive sound and music in Collins' terminology reacts to direct inputs from the player like navigational orders and button presses. Adaptive sound and music, on the other hand, react not to direct player inputs but to game states, which are parameters in the game such as time-ins and time-outs derived from in-game states like the fictive time of day or the avatar's health condition.

In other words, game music must be able to change according to circumstances. This kind of changeable music is elsewhere referred to as *adaptive music* (e.g. Geelen, Tim Van in, M. K. Collins, 2013). I will argue that applying the term *dynamic* to denote what is adaptive or interactive can be potentially confusing - especially in regard to sound where the term already exists as a reference to amplitude changes. If, for instance, one wishes to transfer the term to an interactively manipulable envelope (as I will do in chapter 9) one would end up with the term "dynamic envelope", which in traditional sound production discourse specifically denotes an envelope that controls amplitude, not an envelope that itself can be changed in real time. The distinction made by Collins, however, between adaptive and interactive is useful to describing the potentially very different relationship between the music and the rest of the game program. To avoid misunderstandings caused by the use of the term *dynamic* in these two very different meanings, in the context of the present thesis I will therefore use the terms *direct* and *indirect adaptive music* to denote what Collins refers to as interactive and adaptive dynamic music respectively.

On a side note one might ask, however, if it really is the *music* that is adapting? The word, adaptive, seems to imply that the music *is* something to begin with that is then *changed* - an original musical state that adapts to circumstances. Strictly speaking, adaptive music doesn't exist to begin with. It is being created (or organized[39]) on the basis of a system that changes its underlying properties for musical output by conforming to extra-musical influences. Furthermore, I find it necessary to elaborate shortly on the nature of this musical adaptation - in the name of academic integrity - as I have taken the position throughout the thesis to regard music phenomenologically. From a phenomenological perspective that sees music as a phenomenon of consciousness - by defining it, as I did in chapter 5, as *that,* which is experienced through the *empathetic co-agency* of *musical listening* - the notion of adaptation might have very different and psychological rather than structural implications. If, in this phenomenological conception of music, the music adapts, this adaptation must take place at the level of the empathetic co-agency - that is, integral to the intrinsic shaping and co-enactment of the musical experience. As such, the term adaptation might refer to for instance the schematic assimilation and accommodation associated with the cognitive philosophy of Piaget presented in chapter 5. Here it can be said that the musical experience per se becomes precisely an expression of schematic adaptation occurring in the field of tension between the structure of the sounding music and the musical conditioning of the listener - the listener's musical *history* as Ligeti would put it. The question of adaptive music as either a phenomenon of consciousness or a structural artifact is merely a matter of whether one takes a phenomenological or structural perspective on the ontology of music. For the sake of clarity, I wish to make clear that I am referring to adaptive music in the structural sense here, and that the adaptation taking place occurs at the level of the underlying properties of the music generating system.

Adaptive music in games implies a range of challenges - a simplistic summary of which can be expressed as three primary pre-requisites: *Flexibility*, *variation* and *trans-diegetic coupling* to in-game parameters.

The first of the mentioned challenges is to ensure that the music, while being able to adapt, nonetheless maintains musical compatibility between adjacent sections. I will refer to this capability of the music as *musical flexibility*, an issue I shall return to below in the section on *rhizomatic* adaptive music.

Another challenge is to avoid players getting tired of the music due to long durations of exposure. For certain games the gameplay may be very long. In MMOs such as EVE Online it may be months or even years. This can make extensive repetition and monotony difficult to avoid with a potential overexposure of musical material, which may lead to *listening fatigue* (K. Collins, 2008, pp. 140). Too much repetition (e.g. due to the looping of musical material) can in this way lead to a deterioration of the gaming experience, and irritation in this respect may act as counterproductive to player immersion. Furthermore, musical elements that are easily recognizable such

[39] Adaptive music may range from being purely generative and based on algorithmic realtime composition to being an organization of precomposed material. I will look further into this in the next chapter.

as memorizable melodies and characteristic harmonic progressions may through repetition encourage listening fatigue further. The avoidance of listening fatigue through *variation* is thus an important challenge in regard to the implementation of adaptive music in games.

Adaptive music is traditionally coupled to the gameplay via so-called *triggers* and *cues*. For direct adaptive music such triggers are thus associated with the direct player input. A typical trigger for direct adaptive musical change could be location-based and tied to a certain spatial field that the player may move his or her avatar in and out of. It could also be the player's decision to exit to the main menu of the game that triggered a change in the music. The indirect adaptive music on the other hand is typically associated with so-called RPCs (Run-Time Parameter Controls), which are event- or game state-based cues. RPC cues can be tied to events or states such as night and day, the avatar's physical condition or health, the amount of ammunition left and so on. Another immediate implication of adaptive music for games is therefore the identification of *gameplay parameters* for the music to couple to, as well as the either direct or indirect adaptive trans-diegetic coupling of musical parameters to triggers, cues and RPCs.

7.3.2. COMMON TECHNIQUES OF ADAPTIVE MUSIC

The organization of adaptive musical structure has been commonly approached through four primary techniques. Tim van Geelen identifies these as *branching, layering, transitions* and *generative music*[40] (Geelen, Tim van in, M. K. Collins, 2013). *Branching* refers to the non-linear progressive possibilities of a game as depicted by branches on a tree, where each new branch division represents a choice the player can make – each presenting a potential change in musical expression. The processes for the organization of musical structure associated with this model may involve manually writing an arsenal of musical modules or sequences as well as mapping these modules to the desired branch division on the tree. *Layering* is the process of organizing often pre-composed musical material in vertical layers which can be played simultaneously or separately by adaptively fading them in and out according to gameplay. In situations where the stylistic difference between two

[40] These techniques correspond with the techniques numerous game composers of the industry have presented as their work methods at the Game Music Connect conferences in London for the last three years (Olivier Deriviere, 2013-15) - perhaps with the exception that generative techniques are used less commonly than the other techniques, at least in AAA productions. When asked why this is the case, the composers I have talked to reply that the quality in terms of musicality and sound, which is expected from AAA games today, is not supported by existing generative systems at this time. Nevertheless, it seems clear that the industry is looking towards precisely generative music as a field of interesting possibilities for future development. I look closer at some problems associated with generative music - and suggest solutions - in chapter 8.

successive musical expressions becomes notably abrupt, a variety of composed *transition* modules (stingers) can be implemented to ease the shift. *Generative music* - or algorithmic composition - refers to music which is being composed by the game program as the game is being played, offering a very flexible but also complicated solution to the challenges posed by non-linearity.

I will discuss implications of the branching approach in the context of VDM and sandbox games shortly. Layering is fundamentally a vertically oriented technique (although the musical material in the layers may be in any kind of style and of both horizontal and vertical domination) and is very well suited for VDM for that reason. The need for transition modules may be seen as an expression of the issues arising when music with horizontal dominance is paired with a non-linear medium. Interestingly such modules are often characterized by being vertically dominated in order to break the current musical logic and pave the way for a new - creating a clean slate so to speak. Transitions like this have perhaps less relevance for the implementation of VDM in a game unless one explicitly wishes the musical effect of a transitional musical event. In such a case the transition module performs solely an aesthetic function and not - like it would in a HDM context - a function of hiding abrupt shifts between incompatible musical parts. Generative approaches will be covered in more detail in the next chapter.

7.3.3. ARBORESCENT AND RHIZOMATIC ADAPTIVE MUSIC

Games with a branching, tree-structured (arborescent) narrative allow players to make decisions within a tree-shaped frame of designed possibilities, offering interactivity and entailing non-linearity, but doing so within a static and arguably *linear* frame. A common approach of adaptively addressing such a tree-structure is to design a transition matrix that defines all possible musical transitions as they relate to possible gameplay branchings (K. Collins, 2008, pp. 160).

If a game is not designed according to a tree-structure, however, but is based on, for example, a sandbox principle, then the game design resembles more a *rhizome* structure. In such a case the music might advantageously do the same to be able to couple with the narrative.

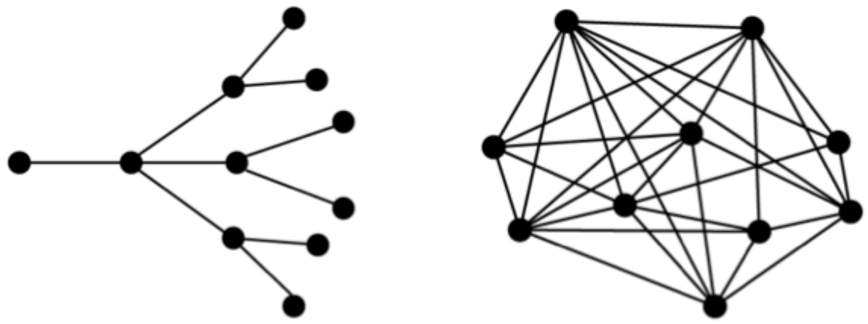


Figure 19 - Simplistic illustration of arborescent and rhizomatic structure

The rhizome as opposed to the tree structure can bear fruit (pardon the pun) also in regard to computer game music. When the narrative in a game is conceived as a branching tree each branching out represents a compositional problem of fitting new music to the old in an aim to synchronize with the action as closely as possible. Adaptive music for such a game structure is therefore often implemented likewise on the basis of a tree-structure that takes into account these different, possible narrative directions through a transition matrix. While this can be effective in many types of games, where the narrative is relatively linear (meaning there are only a few possible paths to the same goal), it poses a serious compositional workload when dealing with a sandbox type MMO such as EVE Online, which has little or no predefined narrative and no designed path from start to end - in fact it might not even *have* a beginning or an end. Such a narrative ideally exhibits no hierarchy, no linearity, no predetermined cause of events, none of the exclusion of narrative possibilities that a rigid tree-structure entails (it's only the stem and the specific branches that are allowed) and a comprehensive transition matrix may be difficult to design.

I need to briefly recapitulate the rhizome structure. Campbell writes about the Deleuzo-Guattarian application of the term in regard to their ”new image of thought”:

> ”The new image of thought is also conceptualized with the vegetal model of the rhizome, and stands in opposition to the traditional image of thought, which is defined as arborescent. In contrast with the hierarchical structured branches found within tree systems, a Deleuzo-Guattarian rhizome has lines which allow the connection of any of its points with any other, and where arborescent systems have ”hierarchical modes of communication and established paths, the rhizome is an a-centred, non-hierarchical, non-signifying system.” (Campbell, 2010, pp. 143)

While a tree-structure can be seen as an expression of directionality in time, and thus an expression of hierarchy and horizontality, the rhizome is timeless and neutral in this respect. A sandbox game cannot be thought of as a tree-structure in the same way, at least not a predetermined tree-structure. It may be seen as arborescent in retrospect: this happened, then this, then this. But, as demonstrated by the arrow of time, irreversible change means that from the point of observation (the now) the number of possible futures is growing, the further ahead one attempts to look - and because of this it is neither possible to completely predict nor undo what has been done. The freedom of the player is greater in a rhizome-structured game design than in a branching tree-structure because there are simply more possibilities offered by the game's rules. This poses a serious challenge in regard to compositional workload because not only does the music have to fit the narrative and be capable of transitioning in a natural way from scenario to scenario, it also has to do so in a way that is not too monotonous and repetitive.

As this is particularly true in terms of narrative (seen as the course of events), when it comes to the gameplay, it is possible to identify trans-diegetic coupling opportunities in most games that can be used in a systematic way. A sandbox battle game may have an almost infinite number of locations to visit, and it may even be possible to travel between locations through teleportation so that transitions between locations may happen rhizomatically. On the other hand, one might identify a general gameplay transition between, for instance, the two states ”non-combat” and ”combat” and couple the music adaptively to these. In any case, such a gameplay-based transition matrix represents a limitation of diversity in comparison with a rhizomatic adaptive approach as some transitional possibilities are left out of the design process.

Genres such as MMOs, sandbox games and any game with a persistent universe thus, and perhaps more than any other game type, put on center stage the music's capability for being *flexible*, and demand techniques for maintaining musical compatibility between adjacent sections. It is here a crucial point that any compatibility issues stem entirely from horizontal dominance. As just discussed for tree-structured narrative of a game's design, so, too, does HDM per definition exhibit a ruleset that narrows down the available directions the music might take at any given moment to encompass only those compatible with the musical logic at hand. This observation in conjunction with VDM's characteristic of *unspecificity* discussed in chapter 4 makes VDM more suitable for sandbox games on a structural level and in regard to the ease of implementation. In this respect it can be said that the more dominated by verticality a given music is, the more flexible it is, due to the fact that fewer compatibility issues will arise at adaptively triggered transitions. As such, the rhizomatic musical structure associated with VDM is also more forgiving than HDM in regard to *open form* - one of ten approaches highlighted by Karen Collins to achieving variation in game music.

7.3.4. MUSICAL VARIATION IN ADAPTIVE MUSIC

Karen Collins lists "ten approaches to variability in game music" (K. Collins, 2008, ch. 8). These are variability through *tempo, pitch, rhythm/meter, volume/dynamics, DSP/timbres, melody, harmony, mixing* as well as two different approaches to variable form, namely *open form* and *branching.*

The need for variation is, as mentioned, a general issue for music in games, especially when gameplay durations are long. In regard to VDM the need is even greater as one of the primary aspects of VDM as a music compositional technique is its fundamental tendency to avoid repetition. In this respect it arguably makes little sense to speak about variation at all as the term implies some sort of periodicity to deviate from. However, the continuous development characteristic of VDM needs musical parameters to function by. In chapter 3 I listed a series of vertical parameters that may serve this purpose. The form-shaping principles of VDM on the horizontal axis are taken up in chapter 9. Based on the above-mentioned approaches to musical variation in games, I will, however, already here begin to open this discussion as some of the musical parameters mentioned by Collins above (such as tempo, rhythm/meter, melody, open form and branching) are highly horizontally oriented. VDM does not ideally involve melody, and as any considerations in regard to branching are largely only necessary due to problems of musical compatibility of a horizontally dominated nature, this aspect also falls outside my immediate interest. I will therefore confine myself here to discussing only tempo, rhythm, meter and open form.

Traditionally, *tempo* can be said to refer to the frequency of fixed musical metric periods (e.g. defined on the basis of quarter-notes as a BPM count - the number of beats per minute). Since in VDM there is per definition little periodicity - in fact not only in regards to meter but to other musical parameters as well - the notion of tempo as such is difficult to apply here. However, VDM equivalents to tempo may be defined in terms of the *average duration and occurrence frequency of spectromorphologies* as well as the *average duration and occurrence frequency of pauses* in-between these spectromophologies. These characteristics may stand as equivalents to also rhythm and meter and can be referred to as defining for the *horizontal meso-density* of a piece of VDM.

In regard to achieving variation through *open form* (i.e. a potentially arbitrary sequencing of musical sections) and *branching* any compatibility difficulties between adjacent musical sections arise due to horizontally dominated musical structure. Arbitrary sequencing of precomposed VDM material, whether presented freely or organized into an arborescent structure, is not a problem in this respect because VDM is per definition devoid of any horizontally dominated musical logic that would have to be maintained through the transition. As far as VDM is concerned, any transition is thus musically viable as long as it is done in a gradual and subtle way in order to avoid unnecessary and potentially unwanted gestalt formations and resulting breaking of the sound-mass. Furthermore, any horizontally dominated musical meso-structures that might reside within the vertically dominated music are in themselves freed from their causal responsibilities - they function as horizontally dominated fragments, detached from any larger causal construct.

An adaptation of Collins' list of variables to suit a VDM context may therefore look as follows:

- Register (pitch)
- Horizontal meso-density (tempo, rhythm and meter)
- Volume/dynamics
- DSP/timbres,
- Harmony
- Layering (mixing)
- Open form
- Branching

7.4. IMMERSION AND AVOIDING COMPETITION FOR PLAYER ATTENTION

In general, music in games must function within the frame of the offered technological possibilities, and for coherency to occur between the game world and the music it must also be adaptive - it must conform to the non-linearity of the computer game medium. Such coherence may have implications to the degree of player *immersion*. As immersion is generally desirable, a main concern, also for music implementation, is to support player immersion and, as a minimum, to not work against it. I will therefore mention, very briefly, a few aspects of immersion here, while making clear that my covering of the subject is peripheral.

Oliver Grau describes the notion of *immersion* as:

> "[…]characterized by diminished critical distance to what is shown and increased emotional involvement in what is happening." (K. Collins, 2008, pp. 133; Grau, 2003, pp. 13)

Immersion is frequently discussed in relation to computer games and virtual reality in academia (e.g., K. Collins, 2008; K. Collins et al., 2014; Ermi & Mäyrä, 2005; Glassner, 2004; Grau, 2003; Slater, 2009) as well as being approached as a common success criterion in the computer game industry.

The term is highly related to the notion of *presence*, which has been described as: *"[...]the perceptual illusion of nonmediation."* (Lombard & Ditton, 1997; in Nordahl & Nilsson, 2014); *"...the feeling of being in an external world"* (Waterworth & Waterworth, 2013); or as Mel Slater has described it:

> ”…the extent to which the unification of simulated sensory data and perceptual processing produces a coherent ‘place’ that you are ‘in’ and in which there may be the potential for you to act” (Slater, 2003, pp. 2).

Additionally, *immersion* and *presence*, are sometimes referred to interchangeably: *”Total immersion is presence”* (Brown & Cairns, 2004, pp. 2).

The notion of presence poses a rabbit hole of existential questions that I will not go into here, however. Rather, I will briefly discuss a few aspects of immersion with relation to computer games. In the context of my focus in this thesis, it suffices to note that the notions of immersion and presence are tied to the quality of the gaming experience and that a high degree of immersion and presence is generally sought after in game development. This poses a demand on the music of a game to help cultivate immersion and presence as best possible - and, equally significant, to not work against it.

Ermi and Mäyrä divide immersion into three categories: *Sensory*, *challenge-based* and *imaginative* immersion. Sensory immersion, in the context of computer games, arrises when the sensory stimuli of the game overpowers the stimuli of the physical real world environment in which the player is situated and thus wins the attention of the player. Immersion through *challenge-based* interaction refers to the way in which the performing of tasks or solving of problems within the gameplay immerses the player in the virtual reality of the game in a balance between challenge and player ability. In *imaginative immersion* the player’s imagination is the key factor for immersion through feelings of empathy with the characters of the game or sheer enjoyment of the fantasy of the game (K. Collins, 2008, pp. 134; Ermi & Mäyrä, 2005).

Based on Ermi and Mäyrä’s three kinds of immersion, it can be said that in some cases, where there is no coherence in terms of synchronicity and narrative expression between music and image (e.g. if the music maintains a peaceful ambient character while the gameplay shifts abruptly from mining an asteroid field into a deadly combat scenario), the music contributes primarily to sensory immersion. In some such cases one may even argue that imaginative and challenge-based immersion is counteracted because auditive stimuli here ignore events and scenarios in the game and therefore present an alternative for the attention of the player rather than a contribution to heightened focus on the interaction with the game’s virtual world. It is an important point that any disturbance of the gameplay, such as unprovoked abrupt transitions (or hard cuts) between musical cues, thus stands for an increase in critical distance and therefore entails weaker immersion, while an elimination of such disturbances may be seen to support immersion[41]. In other words, it has significance to the degree of immersion to what extent the medium

[41] This aspect is related to Michel Chion’s notions of *internal* and *external logic* in regard to the study of sound on film.(Chion, Gorbman, & Murch, 1994, pp. 25-34)

draws attention to itself as a medium. Potentially, the music may sabotage the virtual reality's ability to facilitate the illusion of non-mediation as well as cause an increased critical distance to what is shown and a diminished emotional evolvement in what is happening - thus counteracting *presence* and *immersion* as these are described above.

As mentioned, immersion is a rather hot topic in the computer game industry and is often regarded as a central goal in the development of games making it also an important aspect of academic studies. It seems highly logical, however, that the significance of immersion differs quite a lot between computer game genres according to how important it is for the game to occupy the full attention of the player in order to function successfully. Many relatively simple puzzle games, for example, may be distinguished by a certain lightness of player immersion - such a lightness of immersion being perhaps even a key attraction for the gamer's motivation to play the game. On the other hand, when speaking of, for instance, first person shooter games or other genres where the player in one way or another enacts a virtual identity in the form of an avatar, the immersion itself stands as a crucial precondition for the game's gameplay. (In regard to EVE Online, its status as a roleplaying game in which the player takes on the role of a spaceship pilot means that immersion must be regarded as a central aim.)

As mentioned, whether or not the medium draws attention to itself as a medium has clear implications for the degree of immersion. Although there are exceptions; generally speaking it is critical that the music associated with a game does not take up too much direct attention thereby invading the cognitive space to the expense of other gameplay aspects. In this context the type of attention which is being directed towards the music may be of importance. The Canadian composer and sound theorist, Barry Truax, introduced the notion of "three levels of listening attention", which he calls: *listening-in-search, listening-in-readiness* and *background listening* (Truax, 2001, pp. 21-25). Summarized briefly, when *listening-in-search* the listener has the attention directly focused on the sound, or he or she is searchingly listening for something in particular. This level of listening has the highest degree of focus and attention. *Listening-in-readiness* refers to a level of listening in which the listener is susceptible to significant sounds while not having the attention directed towards any particular sound or the hearing sense altogether. However, the attention of the listener may be caught. In a state of *background listening* the listener has no attention directed towards a given sound, but the sound is nonetheless registered in consciousness and can be remembered if the listener is later asked about it.

In music games such as *Guitar Hero* or *Rock Band*, direct attention on the music is central. In many other cases it may be considered a disadvantage if the music, either through music-structural properties or through its implementation, promotes a state of *listening-in-search* because the primary focus of the player's consciousness is then arguably not directed at interacting with the game. However, trans-diegetic music demands some degree of attention and Truax *listening-in-readiness* may be regarded as a desirable level of listening in this respect. *Background listening* is especially interesting in this context when speaking of non-diegetic music. Music that is meant to facilitate background listing must be composed in such a way as to not draw too much attention to itself. A subtle musical expressivity may thus affect the player on an unconscious level and thereby pose no danger of breaking the immersive illusion of the virtual reality. In the effort to avoid stealing attention from the gameplay or break the illusion it can thus be an advantage if the music is perceived through a state of background listening. This puts certain restraints on the music not to present abrupt musical shifts that can place the listener in a state of listening-in-search or readiness. It should be added that the threshold for how much subtlety is needed, so to speak, to stay within the realm of background listening is relative to the specific gameplay situation the player is occupied with.

7.5. FUNCTIONS OF GAME MUSIC

7.5.1. FUNCTIONS OF MUSIC IN THE COMPUTER GAME MEDIUM

As described in the previous chapter music in audio-visual contexts may lend its communicative properties to other components of the multi-modal phenomenon it is part of. In films, music may have a range of functions which are well suited for this linear audio-visual medium. A range of functions can likewise be listed for computer games. In fact, a key aspect of developing a concept for the implementation of music in games is exactly the choice of function the music should perform. I present in this section a non-comprehensive selection of significant functions of computer game music.

Karen Collins describes six common functions of sound and music in games (K. Collins, 2008, Ch. 7). The functions highlighted by Collins can be categorized as follows:

- Anticipating
- Pointing
- Symbolic (leitmotif)
- Promoting continuity
- Marking formal structure
- Describing the environment

The *anticipating* function is performed when sound or music in a game leads up to a certain event, warns the player or in other ways lets the player know that something is about to happen. A *pointing* function entails that the attention of the player is drawn to a particular aspect of the game. This could for example be an emerging danger that the player needs to react to, or a less immediate component of the narrative that is of significance to the game experience at a particular moment. The *symbolic* function, as described by Collins, refers in the context of music to the traditional technique of leitmotif discussed in the previous chapter. Here the music represents a certain element of the game symbolically. Music may promote a sense of *continuity* when the same musical material is represented in otherwise separate sections of a game, thus defining these sections as part of the same whole. This may be of particular importance if the gameplay is very long like in MMOs. (And, one might add, generally contributes to the aesthetic integrity of a game). Aspects of a game's *formal structure* may be indicated by the music through a marking of changes in scene, game level and narrative such as the beginning or end of particular sections of the game, or as a musical bridge between levels and so on. Additionally, music can function to describe the cultural, physical, social or historical *environment*. Here the music contributes to giving the player a sense of being present in the environment of the game through a musical expressive support and enhancement of the environment's properties. Such a sense of "being present" is thus directly related to the notion of *presence* as this was described earlier in this chapter.

7.5.2. FUNCTIONS OF SOUND IN AN ACOUSTIC ECOLOGY

One academic source for how sounds[42] may function in games in relation to soundscapes is presented through an adaption of the ideas associated with *acoustic ecology*. The research field, *acoustic ecology*, which was formulated in the 1970s by Murray Schafer, is interested in sound as an aspect of our culture and the human relationship to its acoustic environment - or surrounding *soundscape* (Schafer, 1977/1994). Mark Grimshaw and Gareth Schott accommodate the ecological approach to describing the relationship between the player and the virtual environment of first person shooter games. They propose four sound categories to describe sound functions in the game world's virtual acoustic ecology or soundscape (Grimshaw & Schott, 2007). *Choraplasts* are sounds that define a given resonating space or virtual acoustic room in a game. *Topoplasts* sonorously define a specific location in a game. *Chronoplasts* describes sounds that communicate temporal movement - time passing. And *aeonoplasts* are sounds that give connotations to a specific era or historic period.

I include the notions here in relation to non-diegetic *music* despite their original anchor in diegetic and trans-diegetic sounds of a virtual acoustic ecology. I do so to make the point that the *state of the undisclosed sender* and the properties of the *great beyond of the non-diegetic*, in which this sender resides, are not, in my opinion, satisfactory accounted for by an approach to understanding film music and computer game music that takes only a musical perspective. These concepts were conceived in the context of an acoustic ecology, in which the player is a component, and it is thus seemingly problematic to transfer them to non-diegetic music. However, in the discourse of this thesis, which regards music listening as empathetic

[42] I.e. sound, not music. Implications of using this sound theory in a musical context is discussed shortly.

co-agency and the non-diegetic narrative layer as an abstract imaginary ”space” for this music to function in and from, I will argue that the properties of this space may very well be decodable by the same faculties that decode the soundscape of an acoustic ecology. The communicational attributes of the four types of plasts (i.e. resonating space, geography, passing time and era) - which in the soundscape of an acoustic ecology refer to *concrete* properties of a virtual *physical* environment - may thus, likewise, refer to properties of the *abstract* musical artifact and its special abilities for communicating intentions of an undisclosed sender. In other words, a choraplastic function, for example, when performed by non-diegetic music in games, lends its acoustic spatiality - and what connotation this acoustic spatiality might facilitate in the player such as claustrophobia or agoraphobia - to a virtual *meta-physical* environment. This virtual meta-physical environment, I argue, resides in what I, in the previous chapter, referred to as a ”phenomenologically preconditioned imaginary space potential” - the great beyond of the non-diegetic. From here it projects these properties onto non-physical elements of the narrative (such as the space dimension, a ghost, love, envy, confusion etc.) or other components of the multi-modal phenomenon (such as characters, objects and locations). According to this reasoning, music that is experienced as non-diegetic (this includes *covertly* trans-diegetic music) may hold communicational attributes associated with the *choraplastic, topoplastic, chronoplastic* and *aeonoplastic* functions. And it may let these attributes be associated by the player to aspects of the game’s narrative while, and this is the essential point, the music maintains a non-diegetic nature.

Let me give four examples to concretize the use of these terms in regard to non-diegetic music. As with any other sound, also musical sound carries information about an environment’s acoustic properties. For modern productions of popular music there will often be several such spaces simultaneously (e.g. a short ambient reverb on the drums, a long diffuse reverb on the vocals etc.) This acoustic aspect of the music can appropriately be called *choraplastic.* Music that is associated with a certain location is easy to imagine in a computer game where the music can be implemented to accompany the gameplay, when for example a player’s avatar is situated a specific place. Stylistic and instrumental characteristics may also create stereotypical associations to general geographic areas (such as chamber music for the West or sitar music for the East). Here the music is thus *topoplastic*. An example of a *chronoplastic* musical parameter is tempo. Music can, via its tempo, convey a sense of the passing of time as fast, slow or even as having *stopped*. The latter would be the case for music exhibiting *non-cyclic vertical time* (e.g. La Monte Young’s *”Compositions 1960 #7”* discussed in chapter 2). Musical properties of an *aeonoplastic* kind can be exemplified by the use of traits typical for a certain era in regard to musical style or sound (such as baroque music for the 17th century or blues rock for the 1970s).

7.5.3. FUNCTIONS OF FILM MUSIC

Additionally, inspiration for game-musical functions may be taken from film scholars such as Zofia Lissa. Philip Tagg provides an adaption of Lissa's twelve main film-musical functions (Lissa & Tagg, 2015). These are: *Emphasis of movement, emphasis of real sound, representation of location* (encompassing physical and ethnic location such as countries or being underwater as well as social and historical "locations" such as middle-class and medieval), *source music* (referred to elsewhere as diegetic music), *commenting* (such as an audio-visual counterpoint in which the music seemingly contradicts the action of the visuals), *expression of an actor's emotions, basis for audience's emotions, symbol* (through e.g. leitmotifs)*, anticipation of subsequent events* and lastly *enhancement and demarcation of the film's formal structure*.

I wish to suggest here that adaptive music in games in addition to performing trans-diegetic functions specific to the computer game medium may in general also take on musical functions found in the film medium as well as functions that are normally attributed to the film's sound design. Some of the above mentioned functions are duplicates. This applies to the anticipating function, symbolic function and the marking formal structure. A non-comprehensive list of game-musical functions thus includes:

- Anticipating
- Pointing
- Symbolic
- Promoting continuity
- Marking formal structure
- Describing the environment
- Choraplastic
- Topoplastic
- Chronoplastic
- Aeonoplastic
- Emphasis of movement
- Emphasis of real sound
- Representation of location (physical, ethnic, social, historical)
- Source music
- Commenting
- Expression of an actor's emotions
- Basis for audience's emotions

In a non-linear medium, linear music will have difficulties in performing these functions because they - with the exception of the continuity promoting function - all require a degree of synchronicity and thus of adaptability.

7.6. VDM AND THE IMPLICATIONS OF COMPUTER GAME MUSIC

Thus far I have discussed general challenges of computer game music in regard to technological issues, challenges arising from the non-linearity of the medium, techniques and variables of adaptive music, the aim for player immersion as a key concern for game design and the need to avoid competition for player attention, some genre specific issues in regard to particularly very open-form genres such as MMOs and persistent universe sandbox games, as well as functions of game music. It is now possible to cross-reference these general challenges of the computer game medium with the properties of VDM.

Let us first consider the technological aspect. While VDM certainly must abide to the same technological limitations of the sound interface's capabilities, disk space, memory and CPU load as other music-compositional techniques, it has potential challenges of its own in this respect.

The tendency of much of the existing VDM, by composers such as Ligeti, Penderecki and Xenakis, to utilize a very large scale polyphony with many divisi voices - sometimes encompassing individually scored parts for the instruments of an entire symphonic orchestra (like in Ligeti's *"Atmosphêres"* discussed in chapter 4) - requires significant resources in regard to disk space, memory and CPU load if these many voices are to play simultaneously. The *irregularity* of VDM in terms of irregular variations in dynamics, pitch and micro-texture demands further disk space if implemented on a per sample basis, as the sound samples used for playing back the individual notes must be relatively long if deviations in these parameters are not to become periodic with each loop of the sound file. (In appendix A and B I present practical experiments with the construction of such a sound sample library.) These two issues are expressions of the large vertical complexity required by VDM. This latter issue of irregularity might, however, be solved through the implementation of irregularity on a sound engine basis. (Both a sound engine-based and a per sample-based approach will be discussed in the next chapter.) The technological problems of many divisi voices is perhaps more difficult to solve as control is necessary over individual sub-structures (voices or instrument groups) as a means of narration as discussed in the previous chapter. Sample based VDM, based on for instance an acoustic instrumental paradigm with samples of a symphony orchestra, is thus perhaps more problematic in this sense than synthesizer-based VDM, as synthesis requires far less disk space[43].

[43] Wavetable synthesis strictly speaking requires a selection of small waveforms that must be saved to disk. Granular synthesis likewise requires samples, but the size of these may vary significantly and is in any case modest in comparison to the quite substantial disks pace needed to hold e.g. 76 stereo samples of symphonic instruments.

As already discussed at some length VDM has a very high degree of musical flexibility, as limitations in this respect are entirely caused by the properties of horizontal dominance.

VDM's inherent tendency to avoid repetition makes it less prone to listening fatigue caused by this aspect of adaptive music. However, the musical expressions accomplished by applying a vertically dominated composition strategy may itself arguably become tiresome. The constant cognitive perturbation that can arise from listening to VDM must be considered a high risk factor in regard to listening fatigue. Possible solutions include the implementation of a suitable amount of pauses as well as variation in both horizontal and vertical density.

The need to identify gameplay parameters in the form of triggers, cues and RPCs (Run-Time Parameter Controls) to couple musical parameters to is identical for both adaptive HDM and VDM. The musical parameters for achieving variation are not, however. An adaptation of Karen Collins' list of ten approaches to variation in game music resulted in a list comprising: register, horizontal meso-density, volume, DSP/timbre, harmony, layering, open form and branching. Some of these are already discussed in chapter 3 as vertical parameters of the *vertical expression* (see list of vertical parameters in Appendix F).

The kind of musical subtlety associated with VDM's tendency to avoid horizontally separable gestalts by obscuring onsets and termination of tones and spectromorphologies discussed in chapter 4 can be seen as promoting the level of listening attention referred to by Barry Truax as *background listening*. This makes VDM less prone to break with the immersive reality of the game by avoiding to provoke an enhanced critical distance to the game caused by abrupt musical changes.

The rhizome structure of VDM fits well with the rhizome structure of persistent universe sandbox MMOs such as EVE Online. This is already covered to some detail earlier in this chapter, and I will not go further into it here.

In regard to the functions of game music, VDM may easily perform the same tasks as HDM, but may do so in a very specific way due to its unique narrative attributes. For example VDM may perform an *anticipating* function by leading up to an in-game event (this must be done with care not to ruin any deadly surprises through covert trans-diegesis, as the composer of *"Alien: Isolation"* described - see previous chapter); it may *point* to the presence of a ghost (e.g. Simply by the sudden occurrence of VDM. This happens on several occasions in the film, *"The Shining"*); it may be *symbolic* of the mystery of outer space or a VDM substructure, such as the subtle onset of the woodwind section may act as a *vertical leitmotif* for a specific character; it may *promote continuity* by musically tying together the gaming experience of an MMO across gameplay pauses of several days; it may *mark formal structure* by occurring, disappearing, or changing its vertical expression at formally important focal points and so on.

There seems to be nothing to suggest that VDM comes short in regard to any of the functions of film, sound design and games mentioned in this chapter.

7.7. CONCLUSION

The limits to what is technologically possible have moved significantly since the 1970s although they are set by practically the same barriers now as then. These barriers include the quality of the hardware sound interface, disk space, memory and CPU power allocated to music and sound in a game.

The notion of non-linearity may be understood as an absence of a predetermined narrative progression that governs for example the film medium. It entails an indeterminacy in regard to the course of events, which makes synchronization and coherence between the narrative expressions of sound and image a challenge. As a result, linear music implemented in a non-linear medium will neglect most other possible musical functions than that of a musical background tapestry. However, non-linearity is scalable and some game genres are more linear than others.

Game genre categorizations are based on different emphases. These include the visual perspective (e.g. FPS), type of game (such as role-playing, puzzle, persistent universe, multiplayer, single player and so on). These different genres exhibit different degrees of non-linearity. EVE Online, which I present as a case primarily in Appendix B, can be categorized as a sandbox MMORPG with a persistent universe. Games with a persistent universe, sandbox box games and MMOs will often have a particularly high degree of non-linearity and the compositional challenges posed by non-linearity are in turn particularly strong in such games.

To synchronize to an indeterminate course of events, game music must be adaptive music. It can be so in either a direct or an indirect fashion by reacting to direct player inputs or indirect game states respectively. Three primary pre-requisites can be identified for adaptive music in games. These are flexibility, variation and trans-diegetic coupling. The avoidance of listening fatigue, the need for variation as well as the definition of triggers, cues and RPCs to couple the music to are central concerns in the composition and implementation of game music.

Branching, layering, transitions and generative music are common approaches to achieving adaptive music. In the context of VDM, layering is a suitable technique because it is basically a vertical method. Transition modules are to a large extent unnecessary in VDM because the problems that are sought solved by the introduction of such modules are of a horizontally dominated character. Branching, being an expression of an arborescent design structure, entails hierarchy, linearity and limitations in regard to possible transitions. A contrast to the tree-structure is the *rhizome*, in which all points ideally connect to all other points. The rhizome structure may in this respect be seen as an ultimate expression of non-linearity. As such the rhizome is represented in some game designs such as persistent universe, sandbox type games and MMOs whose course of events is too non-linear to be regarded as a tree-structure. I have proposed that music with a rhizomatic structure such as VDM is structurally better suited for such games than HDM because the number of possible transitions between different in-game scenarios may become so large that designing transition matrixes makes little sense.

The long exposure time of many games represents a potentially overwhelming compositional workload in the aim to avoid listening fatigue caused by excessive repetition of material. Achieving variation is therefore a central concern in adaptive computer game music. Various techniques exist for varying musical material. Through an adaptation of Karen Collins' list of ten approaches to game-musical variation, I have identified a list of VDM specific variation parameters including:

- Register
- Horizontal meso-density
- Volume
- DSP/timbre
- Harmony
- Layering
- Open form
- Branching

Immersion and *presence* are generally regarded key goals of game design. *Immersion*, which has been described by, for instance, Oliver Grau as *"[...] characterized by diminished critical distance to what is shown and increased emotional involvement in what is happening"* (K. Collins, 2008, pp. 133; Grau, 2003, pp. 13), may be counteracted by linear music because it does not synchronize to the game's action and therefore presents an alternative to the game's virtual reality rather than supporting it. Additionally, unprovoked abrupt transitions between musical cues stand for an increase in critical distance and thus a weakening of immersion. *Presence*, which is described by, for instance, Lombard and Ditton as *"[...] the perceptual illusion of nonmediation"*, is likewise threatened by anything that will bring attention away from the involvement in the virtual reality and on to the media as a media. Three levels of listening attention, as suggested by Barry Truax, were discussed in the context of player immersion. I suggested that the level of listening attention referred to by Truax as *background listening* is less likely to break with immersion than those of *listening-in-search* and *listening-in-readiness*.

Looking at musical functions in game music, sound design of an acoustic ecology and film music, I synthesized a non-comprehensive list of seventeen functions that may be performed by VDM (in fact by any musical style) in games. This list comprises the following game-musical functions:

- Anticipating
- Pointing
- Symbolic
- Promoting continuity
- Marking formal structure

- Describing the environment
- Choraplastic
- Topoplastic
- Chronoplastic
- Aeonoplastic
- Emphasis of movement
- Emphasis of real sound
- Representation of location (physical, ethnic, social, historical)
- Source music
- Commenting
- Expression of an actor’s emotions
- Basis for audience’s emotions

I argued that VDM may perform any film-musical function and take on functions normally attributed to sound design as well as perform trans-diegetic functions specific to the computer game medium.

VDM is in many respects favorable in regard to the general implications of sound and music in the computer game media because a majority of the issues arising from the meeting between non-linear medium and music as a linear art form come out of the defining factors of horizontal dominance. The unique structural and perceptual attributes of VDM places it in an advantageous position as compared to HDM in regard to all of the above-mentioned game-related implications. One must not overlook, however, the existence of some technological challenges not least associated with VDM’s tendency for very large-scale polyphony and demands for a high degree of complexity in its *vertical expression*.

I have so far discussed some general implications of computer game music, but have yet to address the most fundamental challenge of them all - the actual generation of musical material. This task is in many cases taken care of by a human composer, whose material may be aided by some of the methods for variation and adaptability introduced in the present chapter, so as to occupy longer periods of time without getting boring or tiresome. When gameplay becomes very long, however, as in, for instance, an MMO, the amount of material needed may increase proportionately, and the need for flexibility and variation might benefit from a *generative* approach. This, amongst other things, is the subject matter of the next chapter, which looks into the techniques and challenges of computer-generated music in a computer game context.

CHAPTER 8. COMPUTER-GENERATED GAME MUSIC

From the general challenges of computer game audio and music, it is now time to take a closer look at issues concerning the use of algorithms and generative methods in the composition of game music. In the previous chapter, three primary pre-requisites were identified for adaptive music in games:

- Flexibility
- Variation
- Trans-diegetic coupling

Furthermore, VDM demands a high degree of complexity in its *vertical expression* and must be able to perform a range of the game musical functions already covered. In the present chapter I will investigate methods for meeting these demands through algorithmic composition. While offering great adaptability, which ensures a good potential for flexibility, variation and trans-diegetic couplings, algorithmic composition comes with significant aesthetic drawbacks. This chapter sets out to understand how algorithmic approaches might be developed that are specifically aimed towards adaptive VDM in games, while at the same time they may be designed to balance advantages and disadvantages of computer-generated music.

8.1. INTRODUCTION

This chapter is devoted to establishing a functional and aesthetic contribution to the theoretical foundation for adaptive VDM in games by addressing both advantages and challenges that arise from realizing music via computers through algorithmic techniques. Throughout the chapter it is identified how conventional approaches to automated algorithmic composition, which may be categorized into *transformational* and *generative* algorithms, are not ideal in terms of their ability to produce VDM. In order to build a system specifically with the purpose of composing VDM, a new and specialized approach must therefore be developed.

I present in this chapter a categorization scheme for algorithmic systems that is based on a systems *basic unit of composition*. This is useful because the shorter the unit of composition, the greater becomes the control and adaptability offered by the system, and in turn the greater is the systems' potential for realizing game-musical functions and synchronizing to in-game events. However, with great power comes great responsibility - and for algorithmic composition great challenges in regard to aesthetics. The main aim of this chapter is to identify the most suitable approach for adaptive VDM composition in games based on a compromise between on one side the degree of control and adaptability and on the other side the aesthetic issues of achieving a musically and sonorously convincing output.

The chapter is divided into three main sections, each offering different perspectives of insight into the implications of using computer-generated music in games. The first section concerns itself with general issues of *algorithmic composition* and includes an overview of a number of general aspects of the field including the notions of *mapping, encoding, representation* and *predisposition.* Algorithmic compositional approaches within the categories of *transformational* and *generative* methods are discussed and their potential in games and as basis for VDM is assessed.

The second section of the chapter deals with challenges of using *generative music* in games. These challenges can roughly be divided into two main directions. One is concerned with the degree of adaptability offered by different compositional approaches. The other is concerned with aesthetic challenges associated with music entirely produced within a computer. The focus in this respect is put on challenges associated with *musical coherency, complexity of the sounding musical manifestation* and *basic units of composition.* By examining the relationship between these focus points a framework for assessing the potential of different algorithmic approaches for VDM composition in games is devised.

Based on this framework, which shows a direct relationship between the degree of adaptability and aesthetic challenges, two additional categories of algorithms specifically aimed at computer-generated VDM in games are proposed. These are the subject of the third section of the chapter. The methods of *Vertical Extension-transformational* and *micro-generative* algorithms are here suggested as categories for algorithms that have as their basic unit of composition the *vertical extension* time scale and the *micro* time scale respectively. These additions come with benefits and challenges of their own in terms of adaptability, aesthetics and compositional workload of the system.

On the basis of the included advantages and disadvantages of linear music, transformational and generative algorithms as well as Vertical Extension-transformational and micro-generative algorithms, the chapter concludes by assessing that a transformational approach based on the vertical extension time scale, but with some micro-generative functionality, may be the best suited approach to specialized computer-generated VDM in games.

8.2. ALGORITHMIC COMPOSITION

8.2.1. A SHORT NOTE ON THE CONCEPTS OF AUTOMATED MUSIC GENERATION

The notion of algorithmic composition may be associated with a number of terms such as *generative music* (e.g. Eno, 1996), *procedural music* (e.g. K. Collins, 2009; Wooller, Brown, Miranda, Diederich, & Berry, 2005), and *formalized music* (e.g. Xenakis, 1992). Although not necessarily synonymous, these terms are often used interchangeably to describe the composition of music through automated or rule-based processes - usually facilitated by computers.

Algorithmic composition as denoting the deployment of algorithmic procedures to the creation of musical structures is, however, not necessarily limited to computers. Music composition may generally be considered algorithmic in nature in so far as it applies rules to the creation of music - an aspect that arguably makes the concept of algorithmic composition problematic. What is the alternative, one might ask? After all, music composition that cannot somehow be traced back to a set of rules is difficult to imagine. Even the most rudimentary rhythmical banging together of rocks bases itself on a regular pulse as its underlying system, and the most complex and perceptually chaotic free jazz or VDM maintains the use of some pitches rather than others, some rhythms rather than others.

The facilities for musical rulesets, whether obvious or hidden, may vary. It may take the form of conscious and explicit music theory that is taught at universities and conservatories or devised by individual composers. It might be unconscious and implicit in the form of musical conditioning of a human composer. Or it might be facilitated by a computer on which the ruleset has been programmed. All of these possibilities involve rule-based decisions and some degree of socio-cultural pre-conditioning. The simple point I want to include here is that strictly speaking there is always a ruleset involved in music creation - whether conscious or unconscious, explicit or implicit, followed by a human composer or programmed on a computer.

So what is then the alternative to algorithmic composition? Is not all composition strictly speaking formalizable? I will continue my inquiries into algorithmic, generative, procedural or formalized composition on the basis of a perspective on the terms in which a key factor is the presence of some form of extrinsic automation, whether carried out by a computer or other non-human facility. It is such automated compositional techniques that are the main subject of this chapter.

8.2.2. MAPPING, REPRESENTATION, ENCODING AND PREDISPOSITION

Before discussing different types of music compositional algorithms it is necessary to first mention a few essential terms associated with the field. The first of these is *mapping*. As will be covered shortly, *generative* algorithms may be understood to create new data, new musical material. Traditionally, the algorithms that generate these data stem mostly from areas outside of music why the output of these systems is not inherently musical in character. In order for the algorithm to generate music these data have to be translated or *mapped* to musical parameters. The result of the mapping scheme, which may be unique to each system, can be viewed as a *representation* of the produced data in the form of musical parameters (Nierhaus, 2009, pp. 267). The input of for instance an analytically focused style imitation algorithm is governed by a similar but opposite oriented translation process of *encoding* musical material into a form that can be understood by the algorithm. The term encoding can appropriately be used also to describe the process of translating for example RPCs (Run-time Parameter Controls) of a game such as the time of day into data compatible with a music algorithm.

Mapping and the resulting musical *representation* lead to another important term, namely musical *predisposition.* It is described by Wooller et al. as:

> "…a measure of how well the musical representation can portray the musical ideas and facilitate the compositional processes that are required." (Wooller et al., 2005, pp. 7)

As such, not all mapping schemes lead to a musical representation that is suitable for - or *predisposed* to - all kinds of music. Since mapping, representation and predisposition have crucial influence on the musical output of the generative system, considerations on this design layer are of the utmost significance in regard to expressive potential and aesthetics of the system. A central concern of relevance to this thesis is thus whether or not systems that are predisposed to generating horizontally dominated musical structures have mapping and representation schemes that are adequate for the generation of vertically dominated music. I will return to this question shortly.

8.2.3. GENERATIVE AND TRANSFORMATIONAL ALGORITHMS

Automated algorithmic composition may be based on *generative* algorithms or *transformational* algorithms[44] (Wooller et al., 2005), (K. Collins, 2009). Wooller et al. differentiate between these functions of algorithms by seeing *generative algorithms* as algorithms that create new musical material whereby the amount of data increases. Additionally, generative algorithms tend to increase the musical predisposition of data (i.e. the output data are better suited to represent music than

[44] Wooller et al. include a third category, *analytic algorithms.* Analytic algorithms may be used in the analysis of a musical corpus by extracting features of the music. In doing so these algorithms reduce the amount of data as well as the musical predisposition. Analytic algorithms are left out in my discussion here as they generally concern themselves with the analysis of a musical corpus rather than outputting composed music, which is the focus of my thesis.

the input). *Transformational algorithms* on the other hand transform existing material. As such, transformational algorithms do not significantly influence data size or change the predisposition of the musical representation[45].

In my opinion, a distinction between transformational and generative algorithms is useful from a practical perspective as there is a clear difference between the tasks of recombining existing material and composing new material from scratch, so to speak. And the distinction seems viable also from a perceptual perspective (at least for HDM) because recognizable musical structures may survive in the recombinatorial process and continue to be heard as variations of some original musical structure.

However, from a structural perspective and as part of looking for potential new approaches to algorithmic composition, I see a problem in regarding a growing data size as a defining factor in this distinction. The problem is that in any transformation or recombination of existing musical material new data are in fact created out of the combination. If the system is functioning in real-time (or close to real-time), as would be appropriate for adaptive music in games, these musical data are then played back immediately. It is not meant to be saved but is directly manifested into sounding music. To speak about an increase in data size seems based on regarding ”data” as something prior to the musical representation in the sense mentioned above (i.e. the data from which the mapping to musical parameters is carried out). As such, an increase in data size as a defining factor for the distinction between generative and transformational algorithms only makes complete sense if the system either does not realize the music in real-time, or if the generated musical representation data are stored for other reasons. Whether it is transforming existing data into new data combinations or generating new data entirely by rules, if a system realizes the music immediately, one can speak of new data arising.

Furthermore, it is, strictly speaking, possible to regard both generative and transformational algorithms as fundamentally organizing existing material. The difference lies in the character of the corpus being organized. Generative algorithms may be seen to typically organize twelve notes (the twelve notes of the equal tempered tone system) that each exist in a number of durations and velocities. Transformational algorithms organize musical material of a higher hierarchical order - that is, sound objects and meso-structures are combined into new meso and macro time scale structures.

---

[45] The framework for the comparison of algorithmic music system proposed by Wooller et al. extends beyond the concepts I have included here and e.g. takes into account also ”contextual breadth”, which is concerned with the degree of internal structural causality of the algorithmic music system.

It is thus not unproblematic to distinguish between transformational and generative algorithms through the reasoning proposed by Wooller et al. I agree, however, that the distinction itself may be useful. I suggest regarding transformational and generative algorithms as positions on a scale in which they position themselves not based on the amount of new data created, nor in relation to whether or not they change the musical predisposition, but rather based on the time scale at which they organize musical structures.

As a general categorization framework for adaptive game music I thus suggest that systems for computer-generated game music may be set up according to the degree of adaptability or musical flexibility they offer. This categorization is useful because it gives a direct indication of how well a particular method is suited for the performance of game-musical functions. In this respect, more musical adaptability and flexibility allows for a potentially closer relationship and synchronization between the music, gameplay and game narrative. See table below:

Linear music	**Transformational algorithms**	**Generative algorithms**
Macro basis	Sound object and meso basis	Sound object basis
No adaptability	Some adaptability	High adaptability
Low game-music functional potential	Some game-music functional potential	High game-music functional potential

Table 2 - Showing the relationship between game-musical approach, basic unit of composition, degree of adaptability and potential for performing game-music functionality

There is, in other words, a direct relationship between the time scale of a system's basic compositional building blocks, the musical adaptability it offers and its potential for facilitating the game-musical functions covered in the previous chapter. I will return to this in the section ”basic units of composition”.

As mentioned in the introduction, the above table will see an expansion when brought to encompass approaches that are specialized for the generation of VDM. Before going further into this, however, I will look briefly at a number of transformational and generative techniques as these have been realized traditionally. I do this in order to assess their suitability in regard to computer games and as basis for the generation of VDM.

8.2.3.1 Transformational Algorithms

The methods of adaptive music discussed in the previous chapter, such as branching, layering and transitioning, may be seen as transformational. Karen Collins mentions a range of possibilities for musical transformation used in computer games including *recombinatorial* algorithms (K. Collins, 2009, pp. 9) by which the sequential order of the playback of larger musical cues and sequences is carried out by means of either aleatoric or player input dependent triggering. Additionally, a transformational algorithm may base itself on the parameterized playback of smaller musical segments that are combined by the algorithm following a complex transition matrix to form a multitude of possible musical sequences (K. Collins, 2009, pp. 10). Other transformational possibilities include tempo changes and the addition and subtraction of instrumental layers.

Middleware products such as Audiokinetic's Wwise (a software platform for the implementation of adaptive sound and music in games) offer a range of transformational functionalities on both the vertical (layering) and horizontal (recombinatorial) axes. I have included some insight into Wwise' applicability as a platform for adaptive VDM in the practical experiments in the context of EVE Online presented in Appendix E.

There is no doubt that in the context of games transformational algorithms are highly useful and their use in the industry well established.

8.2.3.2 Generative Algorithms

Algorithmic composition, in the sense that formalizable methods are used to compose music, can be traced all the way back to the year 1025 and the experiments with automatically generating melodies based on text that were carried out by Guido of Arezzo (Nierhaus, 2009, ch. 1). And it is by no means a new development in computer games either. Here it was used to generate music already in the game Otocky (ASCII Corporation) from 1987 (K. Collins, 2009). My interest in the field is not of an historical character, however, and I will not pursue this aspect of algorithmic composition further here. Instead, I am interested in the potential of traditional techniques as basis for computer-generated VDM in games, which is why a number of these techniques are briefly discussed below, including Markov models, generative grammar, chaotic systems, genetic algorithms and cellular automata.

8.2.3.3 Traditional generative techniques

Markov models

Markov chains or Markov models were introduced in 1913 by the Russian mathematician Andrey A. Markov and were originally designed in the context of language processing.

It is well suited for one-dimensional symbol sequences - that is, the determination of the next note on the basis of transition probabilities from preceding notes calculated with reference to an analyzed corpus of musical data. Markov models function as hierarchical systems of states. The higher the order of the Markov model the more prior states are used in the calculation of the next state and the closer to the analyzed corpus the generated output becomes (Nierhaus, 2009, ch. 3). At low orders Markov models typically cause short motifs and phrases to recur with slight variations (Müller, 2015, pp. 69). Its inherent one-dimensionality is well suited for monophonic voice generation but does not in its fundamental form allow vertical structures such as polyphonic layers and harmony (Nierhaus, 2009, ch. 3). Markov models are capable of running as a continuous process, opening possibilities for real-time manipulation of parameters.

While the latter feature makes Markov models somewhat suitable for adaptive game music, its restriction to monophonic voice generation, and its fundamental design to generate causally interconnected motivic and melodic material makes it a highly horizontally focused facility for algorithmic composition.

Generative Grammar

The concept of generative grammar is originally a linguistic model developed by Noam Chomsky in 1957. It received popularity outside of linguistics as well, inspiring for instance Leonard Bernstein's Norton Lectures on "The unanswered question" of music at Harvard University (Bernstein, 1976).

The theory of generative grammar has been used for musical analysis and composition and applied to jazz, traditional European art music and in the field of music ethnology. Like Markov models it facilitates musical material that is context-sensitive and hierarchically organized and as a result of its origin in language processing, generative grammar is (as Markov models) focused on outputting one-dimensional symbolic sequences. Although the ability to generate sequences can be used to produce progressions of chords, vertical structures like polyphonic meso-textures are not within the grasp of this class of algorithmic compositional formalisms - at least in its initial un-extended form. Generative grammar allows for a small number of rewriting rules to create complex musical structures and is an effective and widely used algorithmic compositional formalism in regard to style imitation, musical analysis and genuine composition (Nierhaus, 2009, ch. 4).

Generative grammar is, at the outset, limited by not allowing a continuous musical flow to be output and is confined to finishing its compositional process before outputting. This characteristic makes generative grammar less suitable for real-time game applications where on-the-go manipulations of musical parameters are more appropriate due to the wish for adaptability and synchronicity between the music and other game elements. Its inherent one-dimensionality (monophony) and focus on structural causality and hierarchy makes it horizontally focused and less than ideal for adaptive VDM.

Chaotic systems

The description of the behavior of complex systems via *chaos theory* was very popular in the 1980s where for example the visually appealing fractals of the so-called *Mandelbrot set* triggered discussions also outside of the scientific field. Chaotic systems' capability for *self-similarity* is used to generate fractals. This feature has also been used for algorithmic composition. *Lindenmayer systems*, named after the Hungarian biologist Aristid Lindenmayer, are chaotic systems that simulate the growth processes of plants and, like generative grammars, function based on rewriting rules. Lindenmayer systems successfully produce self-similar structures such as fractals and are effective catalysts for the generation of musical structure in the context of algorithmic music composition when these fractal properties are mapped to musical parameters. The outcome of chaotic systems can be difficult to predict, and changes to the initial conditions of the generative process can have dramatic consequences for the output. The chaotic properties of such systems can offer highly complex and some times musically satisfying internally coherent structures. Some of the chaotic approaches in which it is not possible to intervene once the process has begun, can pose difficulties in realizing precise compositional intensions because much control is given over to the complex behavior of the system (Nierhaus, 2009, ch. 6).

Some chaotic approaches to algorithmic composition are impossible to further control once the initial values are set, thus to some extent limiting their applicability in real-time contexts such as games. Others, such as Lindenmayer systems, do not have these limitation to the same degree. However, self-similarity, which can be seen as the most significant feature of this algorithmic approach, is in the reasoning of this thesis a clear expression of horizontal dominance (i.e. the concept of self-similarity implies repetition and causal development) and is not a desirable capability in the context of VDM.

Genetic Algorithms

The origin of the genetic algorithms used for music generation is associated with the two Americans, John H. Holland and David E. Goldberg and took shape in the 1970s (Eigen, 1973; Holland, 1975). Like chaotic systems, genetic algorithms emulate biology, here as a probability-based search technique modeled on Darwin's theory of evolution. Genetic operations inspired by biological phenomena such as *crossover, selection, survival of the fittest* and *mutation* are carried out in a virtual biological setting creating *generations* of structures based on musical *chromosomes*. A chromosome in this respect can be chord progressions, motifs, phrases or other musical sub-structure. The generated results are evaluated by the system based on defined criteria, which play the role of natural selection, before a new generate-and-test run is initiated. Genetic algorithms tend to create continuously changing music through multiple cycles of production, modification and examination - promoting viable musical outputs in regard to genuine composition, while being less suited for style imitation. It does so either for a specified period of time or until the set fitness criteria have been met. Genetic algorithms tend to produce a multitude of small segments, but have difficulties in creating larger coherent musical structures (Nierhaus, 2009, ch. 7).

Genetic algorithms might be purposely implemented in games based on their ability to output data continuously. And the capability for generating continuously changing musical structure positions genetic algorithms somewhat closer to the structural attributes of VDM. However, the gradual development of material based on horizontally dominated chromosomes reveals that a clear orientation towards a horizontally dominated musical ideal lies at the heart of also this algorithmic approach.

Cellular Automata

More suitable for genuine composition than style imitation, cellular automata, like genetic algorithms, allow for relatively simple initial rules to promote very complex behavior. The groundwork for cellular automata can be traced back to the 1940s and 1950s and is associated with especially Konrad Zuse, Stanislaw Marcin Ulam and John von Neumann. It wasn't until the 1980s, however, that this class of algorithms had its deployment in regard to music composition by, among others, composer and artist Peter Beyls. Cellular automata have their basis in biology in common with chaotic systems and genetic algorithms, but, apart from that, bear more resemblance to Lindenmayer systems in regard to functionality. The concept of music compositional cellular automata is based on the behavior, interaction and state of *cells* in a virtual *cell space* of a varying number of dimensions mapped to musical parameters. Cellular automata do not allow the analysis of a corpus and are thus put to use primarily as facilitators for genuine compositional strategies (Nierhaus, 2009, ch. 8).

The continuous modification of cells, which is an inherent feature of cellular automata, offers the possibility for real-time manipulation of the flow of musical output and strengthens this class of algorithms' potential for implementation in computer games. As is the general trend with the covered algorithmic approaches, cellular automata are also used with the specific purpose of facilitating causal direction and horizontally dominated coherent music.

The tendency of traditional generative designs being clearly oriented towards horizontally dominated structures in their use for music composition applies to also other approaches such as *artificial neural networks*, *transition networks* and *artificial intelligence* that I will not further elaborate upon here.

Summary of traditional generative techniques

Symptomatic for the field of algorithmic composition is an emphasis on style imitation and a tendency to approach music generation through foremost horizontal means. That is, research in the field has mainly been concerned with the generation of horizontal cause-effect relations through pitch, rhythm and harmonic progression rather than vertical parameters such as timbre, instrumentation and texture as well as the horizontal musical development through these parameters.

Seen as a rhizome structure, VDM is per definition not amenable to a generative model. I will here recapitulate the Deleuze quote from chapter 4:

> ”…a rhizome is not amenable to any structural or generative model. It is a stranger to any idea of genetic axis or deep structure.[…] It is our view

> that genetic axis and profound structure are above all infinitely reproducible principles of *tracing.* All of tree logic is a logic of tracing and reproduction.[…] The rhizome is altogether different, a *map not a tracing."* (Deleuze et al., 1987, pp. 12)

The rhizome-like structuring scheme of VDM thus categorically speaks against generative methods as a means for composition.

While some advantages and disadvantages have been mentioned in the above overlook of traditional algorithmic compositional paradigms, in the larger context of hybrid systems and endless possibilities for algorithm expansion it cannot be ruled out, however, that all of the above may be adapted or otherwise deployed in a larger programmed context where the generation of perceptually satisfactory results are output both in regard to implementation in computer games and as a basis for VDM.

The adoption of the aforementioned algorithmic paradigms in music compositional contexts generally witnesses a pursuit of formalisms that can generate self-referential material, thus giving rise to musical meaning through repetition, causal direction and other forms of some perceivable horizontally dominated logic in the musical output. In such cases, all sails are set in an attempt to avoid the merely arbitrary, unpredictable, random and inconsistent - all of which may be considered criteria of *success* to a vertically dominated musical ideal. In regard to musical adaptability, however, algorithmic approaches are highly interesting in the context of games.

8.3. CHALLENGES OF GENERATIVE MUSIC IN GAMES

Putting the tendency of traditional methods to focus on horizontal dominance aside for a moment, having established that the use of a generative algorithm carries the potential of increasing the degree of adaptability and thereby the potential for performing game-musical functions compared to linear music and transformational algorithms, it is now time to look at some of the less flattering aspects of generative music - both inside and outside of the context of games.

While covering also the benefits of what she refers to as *procedural* music in games, Karen Collins discusses a number of disadvantages of this approach to adaptive music (K. Collins, 2009). Collins describes, among other things, how procedural music by often failing to be musically meaningful may have difficulties in immersing the player through emotional induction. It is CPU-expensive, companies often do not have the budget for developing the required complex control logics, and most composers are schooled in linear music and may have difficulties incorporating adaptivity into their compositional process. Furthermore, procedural music is hard to market because audiences don't necessarily understand what it means.

Axel Berndt and Knut Hartmann highlight an inherent trade-off between flexibility and artistic quality in regard to what they refer to as *pre-composed music* (made adaptive through various means) and *generated* music (Berndt & Hartmann, 2007) - roughly corresponding to transformational and generative algorithms respectively. While pre-composed music holds the potential for high artistic value, its adaptiveness is limited. Vice versa, generated music offers an inherent high degree of flexibility but is challenged by a ”weak meta-structural coherency” of its musical output (Berndt & Hartmann, 2007, pp. 144). As a consequence they propose a *hybrid approach* that utilizes the best from both worlds, so to speak, by adapting pre-composed music through various parameters, which I will not go further into here.

Gerhard Nierhaus notes how the typically score-based musical representation of many algorithmic systems has its own limitations:

> ”In most cases, systems of algorithmic composition generate - analogous to the score in occidental music - a symbolic level. However, since beyond this level music is a complex phenomenon of interpretation and reception, this fact points out another limitation of generating musical structure using algorithmic procedures.” (Nierhaus, 2009, pp. 271)

In this quote Nierhaus identifies a for this thesis central problem with computer-generated music: the lack of complexities associated with, for example, human musical interpretation. He seems to exclude, however, the possibility of the computer shaping the musical output of other parameters than those offered by the traditional western score. And not only do algorithmic composition systems traditionally not offer functionality beyond the western score, even the expressive potential of a score-based representation paradigm is generally not explored to full potential. This is likely due to the fact that the audio engines - whether synthesizer- or sampler-based - that are used to turn the musical representation from the algorithmic system into an actual manifestation of sounding music are limited or not exploited fully. And due to the fact that the avant-garde notation techniques of, for instance, Penderecki have not yet fully caught on in the field of computer-generated music to the extent that it can be manifested in an aesthetically viable manner.

That music derived through generative techniques has challenges in the context of games is a case also supported by game composers at the yearly Game Music Connect conference in London the last three years (Olivier Deriviere, 2013-15). There seems to be a general openness amongst the game composers of AAA[46] games present at the conference towards generative techniques, and an interest in the advantages posed by being able to control real-time generation of music in a game context. But as James Hannigan (Composer on *”Harry Potter”, ”Command and*

[46] AAA games refer to games with the highest development budgets in the industry.

Conquer", "The Lord of the Rings", "Theme Park" and *"Warhammer"* game series) points out:

> "…truly musically convincing results have yet to present themselves before an implementation, which goes beyond simple phrase variations, can be implemented on a grand scale. […] It doesn't sound good enough yet." (Hannigan, 2015).

Collins' mentioning of a lacking musical meaningfulness, Berndt and Hartmann's comment on weak meta-structural coherency, Nierhaus' note that score-based representation lacks a layer of interpretation and Hannigan's demand for more musically convincing results all point in the same direction: That generative music has a serious aesthetic challenge - that "it doesn't sound good enough", as Hannigan comments. I will focus my attention here on two primary concerns of this aesthetic inadequacy:

- The generation of *musically coherent* meso-textures on par with that of human composers
- A sufficient degree of *complexity in the sounding musical manifestation*

The next two sections are therefore dedicated to the notions of *musical coherency* and *complexity in the sounding musical manifestation.*

8.3.1. MUSICAL COHERENCY

In regard to musical coherency it is useful to recapitulate briefly the three primary characteristics of VDM, which I identified in chapter 4: to *avoid horizontally separable gestalts;* to *avoid perceivable horizontal regularities;* and to *weaken the perceived bond between spectromorphologies.* When, for instance, Berndt and Hartmann talk of *meta-structural coherency* they are - in the discourse of this thesis' reasoning - primarily talking about maintaining a high degree of *horizontal dominance.* Coherency in this respect means *ensuring* perceivable horizontal regularities and *ensuring* a *strong* bond between spectromorphologies - both of which are only possible if the music presents horizontally separable auditory gestalts.

Musical coherency in VDM is different. It can be said to exist despite the omission of said horizontally dominated musical traits and might be understood instead as a protection of the integrity of the vertical dominance. As covered earlier in the thesis, this may be achieved by way of:

- Abiding to the three primary characteristics of VDM (recapitulated above) throughout the evolution of the generated music
- Aiming for the anticipatory listening position of *perturbation-based anticipation* covered in chapter 5 (by having high musical entropy, no

causal direction, no repetition and by events taking place through irreversible change)

Still, while following these guidelines, an aesthetic challenge remains, also for VDM, in structuring a *temporal* development of vertical expressions - an *irreversibly changing* musical flow - that is capable of maintaining the illusion of a *sender* with whom to *empathetically co-act* (a prerequisite for *musical listening* as I defined it in chapter 5). This is a crucial point in regard to both computer-generated HDM and VDM that I will look further into in chapter 9.

8.3.2. MUSIC-SHAPING STAGES - AND THE COMPLEXITY OF THE SOUNDING MUSICAL MANIFESTATION

When I speak of the *complexity of the sounding musical manifestation* I am referring to the complexity in terms of phrasing, articulation, acoustics and vertical expression associated with the sounding musical expression as heard by the listener.

In the symphonic tradition, sounding music, written and performed by humans, comes into being through a series of music-shaping stages from initial imagined idea over documentation in a score to interpretation by musicians and conductor, perhaps a phonographic recording before it finally arrives at the listener's ear. From here, highly subjective cognitive premises take the musical information and interpret it on the basis of personal conditioning and history (musical history in the sense put forth by Ligeti as discussed in chapter 5). In this process, the music takes many different forms. In the case of an orchestral tradition these different stages might be listed in a simplified manner as follows:

- Musical idea
- Musical score
- Music as a sonic manifestation
- Individual part
- Reconstructed whole
- Acoustic imprint of a specific room
- …and further on to mediation or directly to the human ear.

The *musical idea*[47] here refers to the imagined music as conceived in the mind of the composer, so to speak, and is, importantly, also generated on the basis of culture and history. The written *musical score* shapes the idea by confining it to limits of graphical notation - be it traditional scores or more avant-garde approaches. I have earlier discussed the expansion of the written score through the example of Penderecki in chapter 4. *Music as a sonic manifestation* refers to the physical shaping of the music by specificity of instrument. The *individual part* of the music concerns the shaping of the music by the individual musician. As a *reconstructed whole* the music is shaped by orchestra and conductor as they play it together. From here the music is shaped by the unique *acoustic imprint* of a specific room by picking up spatial characteristics of the venue. The music may further go on to mediation entailing music production processes or flow directly to the human ear. Here the process continues in the human brain implicating subjective as well as generally human cognitive processes associated with the listening experience as discussed in chapter 5.

For fully computer-generated music, which typically realizes its musical representation through sample libraries or synthesis, the music-shaping process is quite different. This is significant because all of the above mentioned stages can be said to introduce varying degrees of indeterminate qualities to the original musical idea - in effect adding complexity to the sounding musical manifestation. When these stages are not present such complexities are lost.

In music where all stages - perhaps after the initial musical idea - happen inside the computer or by other means of extrinsic automation, no indeterminacy is introduced unless it is designed into the algorithm itself. In this way alterations may have been implemented consciously by the programmer, producer or composer in the design of the generative algorithm, but the type of difficult-to-control and highly irregular and nuanced alterations known from acoustic reproduction are to a large extent absent. It therefore represents a design challenge when an absence of most of the abovementioned music-shaping stages deprives the sounding musical manifestation of its ”free” additions of complexity offered by these stages.

It can thus be said that the more adaptive control a system offers the lower is the ”free” additions to the complexity of the sounding musical manifestation offered by music-shaping stages. The relationship between a system’s adaptability and the complexity of the sounding musical manifestation offered by the system is taken up shortly. *Musical adaptability, coherency* and the *complexity of the sounding musical*

[47] The presented list of music-shaping stages is, admittedly, categorical. At least it bases itself on the premise that there is an ”initial musical idea” that is conceived first and which then undergoes change via the different stages. The musical idea is here seen as separate from the musical expression. It is also possible to regard the entirety of the musical shaping process as a systemic generation of a musical idea. The musical expression *is* then the idea, one could say.

manifestation are key aspects of algorithmic composition in games and it seems evident that generative music in games must perform better in terms of aesthetics if it is to gain additional ground in the computer game industry.

8.3.3. BASIC UNIT OF COMPOSITION

As mentioned in the previous chapter, the adaptive organization of musical structure for games has been commonly approached through four primary techniques identified by Tim van Geelen as *branching, layering, transitions* and *generative music* (Geelen, Tim van in, M. K. Collins, 2013). In the following paragraphs, recalling the musical *macro, meso, sound object* and *micro* time scales proposed by Curtis Roads as well as the *vertical extension* time scale that I proposed in chapter 2, I will discuss how each of these approaches to adaptive music are musically predisposed to performing the organization of musical structure based on building blocks that belong to different musical time scales.

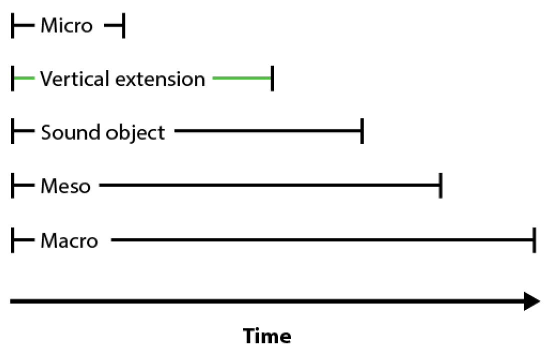


Figure 20 - Musical time scales: Micro; Vertical Extension; Sound Object; Meso; and Macro

In other words, branching, layering, transitions and generative music utilize different *basic units of composition*. This is significant because it may be said that the lower the level of the basic unit of composition the more flexible the musical system can be in an interactive context. Looking at these techniques from the perspective of musical time scales also highlights a shortcoming in traditional generative and adaptive techniques, namely their typical inability to enter the micro time scale which is so crucial to the formation of the *vertical expression* - and thus to the composition of VDM.

As I highlighted in the previous chapter, the processes for the organization of musical structure associated with a *branching* technique typically involves manually composing an arsenal of musical modules or sequences in order to map these modules to the desired branch division on a tree-structured game narrative. The musical progression is then produced by the cause of events in the game as these are tied to a branching ruleset. The basic unit of composition of a branching organizational paradigm is thus typically of a meso level time scale (i.e. characterized by groupings of sound objects such as notes into phrases, sequences, melodies, counterpoint and themes) - or it may be of a macro time scale if one whole linear musical piece accompanies a single branch.

A *layering* technique for accomplishing adaptive music in games refers to the process of organizing typically pre-composed musical material in vertical layers, which can be played simultaneously or separately by fading them in and out in accordance with the gameplay. Music organized through layering in this way is likewise based on the meso level time scale as its basic unit of composition since it is typically meso-structures that are being layered (e.g. a drumbeat, a chord progression or a melody).

Transitions, often organized in a transition matrix, are used to ease difficult shifts between musical segments. These transitions can be meso level musical structures of a bar or a few bars in duration. But they can also, in the case of ”stingers”, be a single intercalated spectromorphology that breaks with the previous musical logic to make way for a new - thus operating on the sound object time scale.

Generally speaking, the traditional algorithmic techniques of *generative music* covered above can be said to organize musical structure of the sound object time scale into higher order meso and macro level musical structures. As the output musical representation of these techniques are limited to primarily score-based symbols (through parameters such as pitch, duration and velocity) not only is music composition confined to the formation and organization of spectromorphologies at the sound object time scale, but the formation of these spectromorphologies is primarily limited to that of musical notes. This effectively leaves out many other possible forms of spectromorphologies of the music composition process. Thus, the traditional generative techniques associated with algorithmic composition can be said to have the sound object (the note) as its basic unit of composition. This makes a generative approach to adaptive music potentially a more flexible solution than branching and layering in terms of the degree of control that may be adaptively exercised. This is because generative music in this sense uses the lowest level time scale unit for organizing musical structure. On the other hand, generative music presents the most difficulties in regard to accomplishing musically satisfying results because it excludes a human composer in the traditional sense and because all involved music-shaping stages are performed within the computer as a complete design.

8.3.3.1 VDM and basic units of composition

As was covered in chapter 2 and 3, VDM is based on the constituents of the *vertical expression* (primarily associated with the *vertical extension* and *micro* time scales) as its lowest units of organization. In chapter 3 I defined the vertical expression as: *an emergent hybrid concept of pitch, harmony, timbre and micro-texture.* Although offering control over which pitches and harmonies to be played and to some extent making accessible changes to dynamics, timbre and instrumentation, none of the abovementioned techniques of adaptive music in games go to any serious depth in terms of micro-texture. Furthermore, none of them offer any significant diversity of vertical parameters as part of their basic musical representation. It is thus evident that while VDM is no-doubt possible to produce through branching, layering, transitioning and traditional generative methods, none of these techniques fully utilize the compositional potential for expressivity and flexibility associated with VDM. A generative system built specifically with the purpose of composing VDM must therefore have an even lower-level basic unit of composition than traditional algorithmic composition, namely those of the *vertical extension* and *micro* time scales.

This does not, as will be more evident shortly, imply that such a system entails the same difficulties that arise in traditional algorithmic composition, as a VDM system does not have to adhere to the stringency of perceivable horizontally dominated musical logic - including causal direction, thematic development, chord progressions, rhythmic patterns and their development.

8.4. APPROACHES TO ADAPTIVE VDM IN GAMES

As mentioned, the smaller the basic unit of composition, the more adaptive the music may be (i.e. the more potential for control is offered by the system that facilitates the music). But as control and adaptability increase the less is given for ”free” in terms of the complexity of musical manifestation (because more music-shaping stages have to be designed) and the more challenging becomes the task of obtaining musical coherency (see *table 8-1* above).

Linear music, which offers no adaptability once it is composed, might consist of orchestral recordings of live musicians and thus include all music-shaping stages associated with acoustic music. The basic unit of composition (if one can speak of composition) would here belong to the macro time scale as whole pieces of music are played back. Additionally, musical coherency in linear music is not compromised by any adaptive changes to the musical development. As has been covered in chapter 7, however, linear music in non-linear computer games is problematic in terms of synchronization and the performance of game-musical functions.

Transformational algorithms allow for the adaptive transformation of pre-composed material. Depending on the level of control offered, the basic unit of composition might belong to either the meso- or sound object time scales. As a consequence of the offered adaptability, however, the ”free” complexity of manifestation associated with this approach is compromised to various extents because some of the music-shaping stages have been transferred to the realm of the computer and thus have to be designed. Transformational algorithmically facilitated music therefore implies that any inherent musical coherency is likewise compromised and must be gained through the careful design of the algorithm itself.

For *generative algorithms* the adaptability goes even deeper. Here the basic unit of composition belongs to the sound object time scale. The compromise at the levels of musical coherency and complexity of musical manifestation caused by generative algorithms' depth of control thus represents a serious design task, which has been approached through the traditional methods of generative music mentioned earlier in this chapter - approached, as was evident, with the aim for horizontal dominance.

8.4.1. VDM AND THE CHALLENGES OF GENERATIVE MUSIC IN GAMES

As so far covered, algorithmically derived VDM must have the vertical extension and micro time scales as its basic units of composition. Musical coherency takes a new meaning in VDM and may be seen to refer to the integrity of vertical dominance. What challenges does VDM pose to the complexity of the sounding musical manifestation?

Significant musical structural and perceptual properties survive during the music-shaping stages - not unaltered, but obviously recognizable. Such properties include for instance pitch, harmony, rhythm, dynamics and instrumentation. These can be seen as simplified musical qualities that translate well between stages. They are the information that the traditional production pipeline of music is built around and therefore cultivates as most significant[48]. As such, the traditional orchestral music-shaping pipeline may be seen to have a musical predisposition of its own, which is not necessarily ideal for a radical approach to VDM composition. The meticulously specific notations of Ligeti's micro-polyphonic approach to sound-mass music, discussed in chapter 4, by which the end goal is that of perceptual un-specificity, is a good example of this. Here the idea that no onsets are allowed to happen at the same time is perceptually more significant than the exact timing of individual entries. But the written score representation, which is necessary when dealing with live musicians, demands exact timings to be notated. The same could be said of the notation of musical structures derived from probability calculations as exemplified in chapter 4 by Xenakis. The inaccuracy with which spectral music accomplishes orchestral synthesis also discussed in chapter 4 is yet another example.

To be free of a score-based musical representation necessitated by human players is thus an example of complexity in the sounding musical manifestation actually becoming easier to obtain through computer-generated music.

[48] It should be added that parallel to the written music tradition, there is a vast field of music culture that preserves other levels of musical information (e.g. which performance traditions come with music of different periods, conformities of instrumental practice, and other performance-related traditions).

8.4.2. VERTICAL EXTENSION-TRANSFORMATIONAL AND MICRO-GENERATIVE ALGORITHMS

On the basis of the research of this thesis so far I will suggest that a specialized system for adaptive VDM in games may be developed by two approaches that are necessary additions to the framework setup earlier in the chapter (*table 8-1*):

- Vertical Extension-transformational algorithms (VE-transformational algorithms)
- Micro-generative algorithms

A revised VDM specific table showing the relationship between game-musical approach, basic unit of composition, degree of adaptability, the potential for performing game-music functionality as well as some mentioned disadvantages can thus be set up as follows:

	Linear music	**Transforma-tional algorithms**	**Generative algorithms**	**Micro-transforma-tional algorithms**	**Micro-generative algorithms**
Basic unit of composition	Macro	Meso	Sound object	Vertical extension	Micro
Degree of adaptability, and control	None	Some	High	Very high	Great
Game-music functional potential	Low	Some	High	Very high	Great
Design challenge for achieving musical coherency	None	Some	Significant	**Some**	Great
Design challenge for achieving complexity of sounding musical manifestation	None	Some	Significant	**Some**	**Immense**

Figure 21 - Showing the relationship between: Game-musical approach; basic unit of composition; degree of adaptability; potential for performing game-music functionality; and the design challenge for achieving musical coherency and complexity of sounding musical

8.4.2.1 VE-transformational algorithms

A future VE-transformational algorithm would use pre-composed *vertical expressions* as its basic unit of composition. That is to say, it performs compositional tasks based on the *vertical extension* time scale. In this way several sub-structural vertical expressions may combine in the sounding simultaneity to form increasingly complex higher-level vertical structures. As the basic unit of composition is lower than that of conventional generative and transformational algorithmic methods, a VE-transformational approach offers an even higher degree of adaptability and in turn game-music functional potential.

The discussed un-specificity, which is allowed naturally by not being confined to for example equal tempered notes, allows for easier vertically dominated musical coherency. That is, compared to realizing VDM through a score-based generative paradigm, VE-transformational algorithms would offer an easier avoidance of separable gestalts, regularities and spectromorphological bonding by not having to specify its structure into, precisely, *gestalts* - the note. This is not least because when vertical expressions are faded in and out of each other they automatically tend to avoid separable auditory gestalts because this omits attacks (note onsets). When the musical note is the basic unit of composition (like in a system with a score-based representation) each note has an attack that has to be somehow obscured. As was evident in chapter 4, Ligeti goes to great lengths to do exactly this in both Atmosphêres and Lux Aeterna. As mentioned above, the emancipation from having to represent VDM via a traditional score, which traditionally limits the capabilities of generative approaches in terms of the vertical expression, potentially lessens the design challenge for complexity of sounding musical manifestation as well.

A VE-transformational algorithmic approach has a quite significant drawback, however. Continuous movements via vertical parameters such as pitch (glissandi) and gradual compression of spectral envelope are not possible because pre-composed vertical expressions are static per definition.

Because VDM does not have to confine to a stringent horizontally dominated causalities and when specificity of individual voices is not a necessity, a VE-transformational system for adaptive VDM might paint with a large brush, so to speak, while still offering micro-textural control. The superimposing of micro-texture onto audio samples of pre-composed vertical expressions is one possibility (e.g. by introducing irregularity of pitch and amplitude to a harmony-timbre). The adaptive layering of vertical substructures to form higher-level vertical expressions is another possibility. These possibilities of VE-transformational algorithms are demonstrated in Appendices A, D and E in the prototypes: ”NowEngine” and ”KeynotePlaybackSystem” as well as through experiments with Wwise in EVE Online.

8.4.2.2 Micro-generative algorithms

A micro-generative algorithm would be concerned with composing VDM based on the *micro time scale* as its basic unit of composition and offer radical generative VDM through the manipulation of any vertical parameter (see appendix F). As such it operates with a potential for *great* adaptability and control and in turn a great potential for performing game-musical functions.

However, it represents an *immense* design challenge in terms of achieving complexity of sounding manifestation as every aspect of the music has to be designed - the manifestation of all music-shaping stages and musical structures at all musical time scales have to be performed by the algorithm. Although it is not confined to a score-based representation it poses a *great* design challenge in terms of creating musical coherency for the same reasons. It is important to note in this respect that the more tasks are handed to the algorithm the more likely it is to become overly CPU-expensive. A micro-generative approach must be based on CPU-expensive micro time scale sound generation techniques such as various forms of synthesis including wavetable, granular and concatenative synthesis and seems difficult to implement fully at this time in a form that meets the high aesthetic demands of AAA games.

A quick cost-benefit analysis of the covered advantages and disadvantages in the above table shows that a VE-transformational approach may be the best candidate for a specialized adaptive VDM system in games at current. As mentioned it has, however, a particular disadvantage that is quite serious in a VDM context, namely its inability to encompass continuous developments in vertical parameters such as pitch (glissandi) or a gradual compression of the spectral envelope because the pre-composed vertical expressions on which it is based are *static* per definition (see chapter 2). For VDM, which must develop through *irreversible change*, this is unacceptable.

The idea arises to have a system based primarily on a VE-transformational algorithm but with specific micro-generative capabilities to cover limitations such as the inability for continuous movement through some vertical parameters.

8.5. CONCLUSION

Adaptive control is important in game music. Greater control and adaptability in this sense mean greater musical flexibility and in turn more possibilities for ways in which the music may adapt to the gameplay. As a result the greater is the potential for the music to perform various game-musical functions (e.g. anticipating, describing the environment, chronoplastic or emphasis on movement). Furthermore, control over the vertical extension and micro time scales opens for more possibilities of performing several functions at a time than when having adaptive control over merely the sound object, meso and macro time scales. This is because the music can be generatively constituted by a higher resolution of sub-structures whereby the musical expression may incorporate a collective of music-gameplay couplings between musical substructures and in-game narrative elements - all sounding at the same time. For systems with a lower resolution of control (caused by a higher-level basic unit of composition) such possibilities are proportionately more limited. It may be said, therefore, that there is a direct relationship between the basic unit of composition with which a system operates and the degree of potential for the ever important narrative interconnectedness between the gameplay and a game's music.

Not all mapping schemes lead to a musical representation that is suitable for - or *predisposed* to - VDM. A central concern of relevance to this thesis is thus whether or not systems that are predisposed to generating horizontally dominated musical structures have mapping and representation schemes that are adequate for the generation of vertically dominated music.

A distinction between transformational and generative algorithms is useful from a practical and perceptual perspective. But from a structural perspective this categorization scheme is more troublesome as the defining factors of the two categories are strictly speaking interchangeable. Both may be seen to create new data and both may be seen to organize existing data.

I have suggested in this chapter a different scheme, which is based on the time scale at which the organization of musical data is taking place. In this scheme transformational and generative algorithms are placed in the larger context of a scale that ranges from linear music over transformational and generative algorithms to VE-transformational and micro-generative algorithms based on their basic unit of composition.

Traditional techniques for generative music emphasize style imitation and share a very strong tendency to approach music generation through foremost horizontally dominated means. Much effort is made in these approaches to avoid the perceptually arbitrary, unpredictable and random that may be considered criteria of success within a vertically dominated musical ideal.

Transformational algorithms are well represented in adaptive game music. Seen in the perspective of this chapter’s reasoning this may be due to a good balance between adaptability and aesthetic quality.

Generative algorithms are largely absent in AAA games despite their advantages in terms of adaptability and thus potential for synchronicity and narrative coupling to the gameplay. This may be explained by a number of disadvantages held by this category of algorithms. Some of these challenges are connected to the resources necessary to make such a system. Others are aesthetic in nature. I have focused on mainly two aesthetic challenges associated with generative approaches to adaptive game music, namely the questions of *musical coherency* and the *complexity of the sounding musical manifestation*. As the traditional methods of generative music composition are designed for achieving horizontal dominance, their properties in regard to *mapping* and *representation* are insufficient for being directly applied to the generation of vertically dominated music. Additionally, by not offering access to the vertical extension or micro time scale in terms of their *basic unit of composition*, these techniques are not, to any serious extent, capable of generating the most crucial aspect of VDM, the *vertical expression*. Furthermore, the rhizome-like structure of VDM speaks against a generative method as rhizomes per definition are not amenable to a generative model.

In regard to VDM, *musical coherency* may be understood as an expression of the integrity of vertical dominance by 1) abiding to the three primary characteristics of VDM (recapitulated above) throughout the evolution of the generated music, and 2) aiming for the anticipatory listening position of *perturbation-based anticipation.* However, a significant aspect of the musical aesthetic, also for VDM, is maintaining the illusion of a *sender* with whom to *empathetically co-act* - a task that is perhaps generative music's most difficult to achieve convincingly.

The *music-shaping stages* differ greatly between the production pipeline of traditional acoustic music and computer-generated music. The free additions of *complexity* to the *sounding musical manifestation* offered by acoustic music are largely absent in computer-generated music. This represents a loss of complexity that has to be regained, so to speak, through design. The nature of computer-generated music is such that its music-shaping stages have to all be designed. This means that computer-generated music with a high degree of control lacks the "free" additions of musical complexity offered by for instance acoustic orchestral music. The result is a lower degree of complexity in the sounding musical manifestation unless such complexities are included as functionality in the design of the generative system. For VDM, in which the complexity of the sounding now - the vertical expression - is of utmost importance, this challenge is particularly serious.

A specialized system for adaptive VDM must have the *vertical expression* as its lowest unit of organization in order to have a fitting musical predisposition. The vertical expression is associated with the *vertical extension* and *micro* time scales. Of the compared approaches VDM thus demands the most adaptive and therefore also most aesthetically challenging methods in terms of musical coherency and complexity of the sounding musical manifestation: *VE-transformational* and *micro-generative* algorithms. However, since traditional VDM techniques have been historically associated with what I have previously referred to as *un-specificity* and since this un-specificity in entirely computer-generated music does not have to be specified for human players in the form of a score, this, in fact, represents an example of complexity in the sounding musical manifestation actually becoming easier.

A VE-transformational algorithmic approach to adaptive game music offers a very high degree of adaptability and control over the musical flow and thus carries a very high potential for performing functions of game music in synch with the gameplay. The design challenge of achieving musical coherency is relatively low in the context of VDM. This is primarily because the score-based representation, which is rather limited in vertical terms, is left behind in favor of a focus on the vertical expression as a basic unit of composition. Also the design task of achieving a high degree of complexity of the sounding musical manifestation is comparably low in relation to both traditional generative techniques and micro-generative methods. A main reason for this is that when vertical expressions are faded in and out of each other in the name of avoiding separable auditory gestalts (one of the three characteristics of VDM), there are really no attacks - and thus, the separability of auditory gestalts is greatly weakened[49]. In a score-based representation each note has an attack that has to be obscured. Ligeti goes to great lengths to do exactly this. However, a VE-transformational approach is not capable of facilitating certain continuous movements such as glissandi. This is because the ”modular” design scheme of organizing static vertical expressions, while possible to for example fade smoothly in and out in regards to amplitude, does not allow for micro time scale control. This particular problem might be solved by including in the algorithm specific micro-generative functionalities where the VE-transformational technique falls short. Another approach to introducing gradual development of vertical parameters is to superimpose such developments through audio effects such as filtering, pitch shifting or amplitude modulation. Not all vertical parameters are easy or possible at this time to superimpose, however. Gradually changing the spectral spread of overtones in a large number of pre-produced vertical expression, for example, would be unacceptably CPU-expensive.

It is evident on the basis of the present chapter that a VE-transformational approach with limited micro-generative capabilities that will enable it to compose an irreversibly changing musical flow through most vertical parameters is well suited both for game music and as a specialized system for composing VDM.

It is also evident that the one factor that limits transformational and generative methods the most in terms of VDM is their inability to enter the vertical extension and micro time scales. The result of this inability is an aesthetic inadequacy in terms of the composed vertical expressions in such systems. This means that although transformational and generative techniques for adaptive game music are not ideal at the outset, such systems may be expanded to perform better as basis for VDM by introducing control over the vertical extension and micro time scales. This particular aspect of expanding a generative system to access the vertical extension and micro time scales is presented in the form of practical experiments in appendices A and B.

The task of obtaining a high degree of complexity in the sounding musical manifestation is not to be taken lightly, especially in VDM. This challenge is not least concerned with the introduction of *irregularity* on several musical levels. This

[49] Additionally, it is in the attack or onset of spectromorphologies that the regularities of computer-generated music become most obvious. This is briefly discussed from a practical perspective in Appendix B.

question of irregularity is a significant aspect of the experiments presented in Appendix B.

An aspect of VDM composition, which I have not yet covered, is the question of horizontal developments. While having identified that a VE-transformational approach with limited micro-generative capabilities gives access to the most crucial time scales of VDM, the irreversible change with which material at these time scales must evolve is the subject matter of the next chapter. Here an overview of horizontal strategies that might be associated with a system for adaptive VDM in games is provided.

CHAPTER 9. HORIZONTAL STRATEGIES FOR COMPUTER-GENERATED VDM

I have found it necessary to focus heavily on implications of the vertical dimension in this thesis - not least because this is what VDM is dominated by. But a highly important question remains: How may the vertical expressions of VDM be developed through time to become a flow of sounding music?

9.1. INTRODUCTION

Based on the theoretical ground covered through the thesis, this chapter proposes some basic principles that offer possibilities for horizontal development of VDM in games. The principles suggested here are not to be understood as an exhaustive investigation of how VDM might evolve over time. However, I have sought to present a framework that has a sufficiently broad applicability to encompass a wide range of horizontal development schemes.

The enquiries of this chapter are focused on four main topics of VDM horizontality in a game context: the relationship between time scales and horizontal development in VDM; the question of VDM's musical representation; the question of this representation's dependency on gameplay data; and lastly, three methodological approaches to horizontal development in adaptive VDM are discussed.

Firstly, in order to maintain vertical dominance horizontal development of VDM must comply with the structural characteristics of VDM identified earlier in the thesis: namely to *avoid horizontally separable gestalts,* to *avoid perceivable horizontal regularities*, and to *weaken the perceived bond between spectromorphologies*. I suggest that these characteristics, when conceived as proactive compositional principles, do not necessarily have to be followed with equal stringency at all time scales. An overview of the implications of horizontal development based on these compositional principles at the micro, vertical extension, meso, macro and supra time scales is presented.

Secondly, this chapter deals with the question of musical representation in VDM. As I identified in the previous chapter, a score-based representation is not ideal for computer-generated VDM in games, but what then is? I suggest in this chapter that the concept of *envelopes* may be a viable meta-structure for horizontal development of vertical parameters. I propose that such envelopes for computer-generated music in games may be *static* or *adaptive* - and that both of these envelope variants may be multidimensional in order for a single meta-structure to change several parameters.

I further propose that envelopes can be used as a meta-structure also for the superimposing of vertical extension time scale and micro time scale *irregularity* on musical material stemming from a score-based generative algorithm such as CALMUS[50]. As such, envelopes may be used both for *composing* VDM and for *superimposing* micro-textural variation and irregularity.

Thirdly, I will discuss briefly in this chapter the question of game music's dependency on gameplay. The horizontal development of VDM in games, being adaptive, is to some extent following data coming from the game either directly or indirectly (by reacting to direct player inputs or indirect game states respectively as discussed in chapter 7). But although it is following these data adaptively, adaptive game music is most often not entirely controlled or computer-generated on the basis of adaptation to game events. As such, one may speak of *gameplay-dependent* and *gameplay-independent* horizontal development. I make the case that it is crucial to the design of the developmental concept of VDM in games to take into account the combination of gameplay-dependent and gameplay-independent development as co-constituents of the horizontal strategy.

Finally, as a methodological frame for mapping envelopes to musical parameters, VDM may evolve over time based on mainly three methods: *deterministically; indeterministically;* and *stochastically* derived horizontal development. A short discussion of the implications of each of these methods is presented.

9.2. HORIZONTALITY AND MUSICAL TIME SCALES

In chapter 8 I developed a categorization scheme for game music approaches based on the time scales of the basic units of composition. Also the horizontal organization in VDM may be categorized according to time scales. This is useful because the maintenance of the integrity of vertical dominance relies on different rules at different time scales. The reasons for this will become apparent in the following pages.

The distinction between time scales, however, is more difficult in VDM than in HDM because the borders between micro, vertical extension, sound object and meso level structures are per definition blurred by VDM's tendency for smooth, continuous and gradual horizontal development. Spectromorphologies are not necessarily clearly distinguishable. As identified in chapter 4, one of VDM's three primary characteristics is exactly to avoid horizontally separable gestalts. Nonetheless, from a structural perspective (i.e. from the perspective of *composing* VDM rather than perceiving it) a time scale categorization may be useful to provide

[50] Appendix B presents CALMUS and EVE Online and my collaboration with the developers of both.

the necessary concrete guidance in terms of how an adaptive VDM compositional system should organize sound over time at these different scales.

9.2.1. VDM HORIZONTALITY AT THE MICRO- AND VERTICAL EXTENSION TIME SCALES

Compared with all other time scales, the horizontal organization within a coherent *vertical expression* - that is, within the duration of what I have called the *vertical extension* (encompassing what Curtis Roads calls the micro time scale, see chapter 2) - has the strictest obligation to avoid pattern recognition and the formation of separate gestalts. As discussed earlier, the vertical extension is not a fixed duration but may be defined as a period of time during which the listener experiences *auditory verticality*. This relationship was described in the definition provided for auditory verticality in chapter 2:

> **A perceptual phenomenon, which is facilitated by an illusion of ”auditory stasis” caused by structural properties of a ”vertical expression” and occurring within a temporal window of a ”vertical extension.”**

The duration of this temporal extension of verticality is dependent on structural properties at the micro time scale in the vertical expression occupying it. As soon as these structural properties clearly facilitate the formation of new temporally truncated gestalts (e.g. by exhibiting a pulse, perceivable repetition or by changing abruptly), a new vertical expression can be said to have taken over - and the horizontal dimension has gained ground in the battle for domination. Perceptually, at this scale the vertical-horizontal 2-dimensionality is reduced to an auditory *singularity* and the properties of horizontal organization cannot be distinguished from those of vertical organization.

In terms of narrative use in games, maintaining a long and static vertical expression in this way might serve for example a *chronoplastic* game-musical function, one of time standing still. It might be used *symbolically* - tied to a particular event or character, or it could *emphasize movement* - in this case emphasizing that movement has stopped.

9.2.2. VDM HORIZONTALITY AT THE SOUND OBJECT TIME SCALE

Broadening the perspective now one order higher in the time scale hierarchy we arrive at the *sound object* time scale. Horizontal development at this scale concerns the evolution of vertical expressivity internal to the sound object. In order for music to be dominated by vertical expressivity, horizontal development at the sound object time scale must, as is also the case at the vertical extension time scale, avoid structures that will cause the listener to form separate sound object gestalts - thus destroying the sound objects coherency (it would then no longer be one, but two or several sound objects). At this scale, however, the horizontal dimension is both phenomenologically and structurally separable from the vertical - allowing horizontal evolution of the vertical expression to take place inside the coherent gestalt of the sound object. The sound object's integrity is maintained as long as this development is gradual, smooth and continuous - as long as it contains no perceptually separable individual events.

9.2.3. VDM HORIZONTALITY AT THE MESO TIME SCALE

Horizontal development at the meso time scale refers to the sequential organization of sound objects into higher-level meso-structures. At this time scale, which HDM equivalent contains melodies, motifs, thematic development and harmonic progressions, separate gestalts formed by the perception of separate sound objects may exist within the meso structure they collectively constitute. Depending on the desired degree of vertical dominance, the horizontal organization can tie the constitutive sound objects together in a continuous sound-mass that avoids pauses. According to the reasoning of chapter 5 this will encourage an *anticipatory listening position* far towards *perturbation-based anticipation.* VDM might also be composed in such a way that spectromorphologies become clearly perceivable during the evolution of a meso-structure. (Importantly, this does not eliminate the vertical dimension's dominance, but merely decreases it. A significant point in regard to VDM's three primary characteristics is that they are precisely *characteristics*. Any algorithm, however, that sets out to compose VDM must be able to comply in as extreme a manner as possible with these characteristics.) Horizontal development at the meso time scale must in this way *tend* towards perturbation-based anticipation by being unpredictable, highly entropic, containing no perceivable repetition or causal direction and by taking place through *irreversible change* (i.e. by being unpredictable in terms of the development's end goal). Furthermore, spectromorphological bonding must be at a minimum.

9.2.4. VDM HORIZONTALITY AT THE MACRO TIME SCALE

The sequential organization of meso-structures takes place at the macro time scale. Here organization may include the separation of sound object- and meso-structures by pauses of various duration - silences in the non-diegetic narrative layer so to speak - or changes may occur in the flow of VDM that are sufficiently abrupt to cause obvious separations between musical substructures. This is not a problem on the macro time scale level as long as these horizontal formations do not violate the rules associated with perturbation-based musical anticipation by exhibiting obvious patterns, predictability or repetition.

In games, the macro time scale may involve musical constructs delimited by game levels or the period in which the player continuously plays the game. In some cases, depending on the specific game design, it may be the entire duration of a game. Therefore the compositional process associated with macro-structure demands a strategy suitable for this time scale, and design considerations include the overarching game aesthetic and narrative of the game level or entire game.

9.2.5. VDM HORIZONTALITY AT THE SUPRA TIME SCALE

In a computer game context the supra time scale may denote the time span of an entire game or game series. It may include the entirety of a gamer's engagement with an MMO such as World of Warcraft or EVE Online. Compositional strategies at this time scale may advantageously go hand in hand with the overall narrative arch of the game if such one exists. It is at this time scale that the rules of VDM can be said to matter the least. Certain functions of music, such as creating a sense of continuity or a syntactical function in which the music delimits a certain section, level or chapter of the game by signaling the end of one part and the beginning of a new, are arguably of more importance at this time scale than the maintenance of irregularity, absence of repetition and so on. The lesser importance of committing to the underlying structural and perceptual principles that govern VDM is caused by this time scale's detachment from the immediate musical listening experience.

9.3. REPRESENTATION OF VDM

Having established in chapter 8 that a VE-transformational algorithm with limited micro-generative functionality is a well-suited method for adaptive VDM, one may ask: How should this music be *represented* by the compositional algorithm when not through the basic parameters of musical notes (pitch, duration, velocity)?

9.3.1. MULTI-DIMENSIONAL META-STRUCTURES FOR HORIZONTAL DEVELOPMENT

A musical representation that facilitates horizontal development on one or several vertical parameters can be regarded as constituted by abstract *meta-structures.* Such meta-structures can be understood as general vectors or curves (rather than specific pitches, intervals and velocities) and their shape can be superimposed on any vertical parameter (e.g. pitch, amplitude modulation, spectral spread, register range, temperament etc.). An example of a horizontal meta-structure can be found in the automation functionality available in most computer-based DAWs, such as Pro Tools, Cubase, Logic and Ableton Live, in which parameter changes can be animated over time. Automation in this regard is usually represented graphically as a line or graph with a series of user-generated points that represent changes in the graph's direction as time passes.

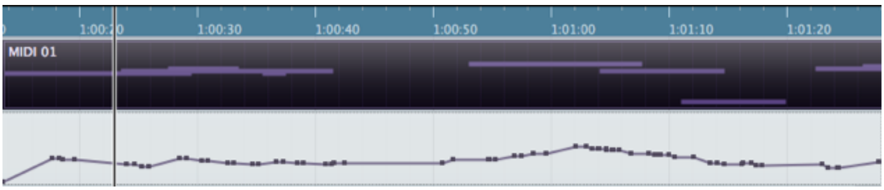


Figure 22 - Automations from the Cubase arrangement of "DistanceMusic1" presented in Appendix C

Such automation typically represents a 2-dimensionality - one parameter is changed over time. In regard to VDM composition where a number of vertical musical parameters can be made to change simultaneously one may speak of a multi-dimensional meta-structural representation - or multi-dimensional *envelopes*.

9.3.2. ENVELOPES

The kind of gradual horizontal development that characterizes VDM can be conceived of as *envelopes*. An envelope in the mathematical sense refers to a curve or curved shape formed by *”a trace of a point whose direction of motion changes”* (TheFreeDictionary, 2016). In the world of sound synthesis, envelope generators are used to control the horizontal development of, for example, the amplitude and filter cut-off frequency of the synthetic tone often via an ADSR[51] envelope framework (Attack, Decay, Sustain, Release). Such usage could very well be included in a system for generating adaptive VDM.

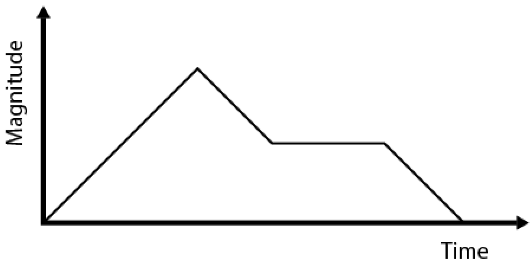


Figure 23 - Illustration of basic ADSR envelope

But the horizontal development of VDM is multi-dimensional, featuring one dimension for each variable vertical parameter. Envelopes for the control of VDM, if they are to take full advantage of the expressive potential of several vertical parameters simultaneously, must therefore also be multi-dimensional. I deal with this further below.

[51] Several other envelope shapes exist as standard on e.g. analog synthesizers - including AHR (Attack, Hold, Release) and AR (Attack Release).

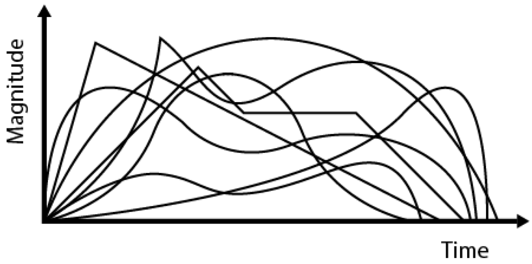


Figure 24 - Illustration of multi-dimensional envelope

Envelopes are effective form-shaping tools that can be deployed on vertical parameters. Any accessible data stream can potentially be used as input to define such an envelope. In a computer game this could be RPC data coming from the game engine allowing the game states to define aspects of the vertically dominated musical flow.

9.3.2.1 Static and adaptive envelopes

Envelopes can be implemented statically or adaptively. A *static envelope* represents a fixed shape, a fixed meta-structure, which can be followed by some vertical musical parameter over time. The shape of *adaptive envelopes,* on the other hand, is generated in real-time (based on for instance in-game RPCs) and can only be overviewed in retrospect. Its shape is generated and followed simultaneously.

The number of possible static envelope shapes is infinite. A system for the generation of adaptive VDM should therefore ideally present an openness towards all these possibilities by allowing the user to build reservoirs of freely determined envelope shapes by, for instance, drawing them in a coordinate system rather than limit itself to a fixed number of predefined shapes. (This is being implemented in CALMUS as a consequence of my involvement in its development. See Appendix B.) For adaptive envelope shaping such an openness could be represented by allowing the user to map in-game data to vertical musical parameters.

A multidimensional envelope - whether static or adaptive - with multiple envelope generating inputs from the game as well as multiple envelope-following vertical parameters may serve as a musical representation better suited for computer-generated VDM than a score-based representation, as it allows for precise control over the kind of gradual and continuous movements associated with *irreversible change* (on vertical parameters where this is technically possible).

9.3.2.2 Envelopes and score-based representation

The use of envelopes as musical representation is well suited for VDM because of this music's smooth, unstriated and continuous qualities described in this thesis. However, the use of score-based representation should not be entirely excluded. VDM is surely possible to achieve through such means - after all it is through a score-based representation that VDM was first conceived. Although composers such as Ligeti and Penderecki felt it necessary to expand the expressive vocabulary of the written score, as we saw in chapter 4, these pioneers of the style wrote music by notating pitches, note durations and dynamics.

The combination of the two approaches opens up aesthetic as well as functional possibilities in situations where, for example, a certain game state is associated with a score-based representation while another game state is mapped to an envelope-based representation. In such cases each of the two representational paradigms constitutes a musical effect each with its own unique musical expressive potential that can be exploited narratively. Moreover, envelopes and score-based representations can be used simultaneously. This is effective in a situation where micro time scale control is superimposed via envelopes on a score-based musical representation coming from a generative system.

An example of this use of envelopes as a means for enriching the *complexity of the sounding musical manifestation* is included in the section of Appendix B where practical experiments with VDM in the generative system, CALMUS, are presented.

9.3.3. COMPOSING AND SUPERIMPOSING ENVELOPES

Two different functionalities are thus possible for envelopes in this respect. One is concerned with *composition* of VDM through horizontal manipulation of vertical musical parameters based on, for example, RPCs. The other deals with *superimposing* micro-textural irregularities on a stream of musical material. In the latter example envelopes further shape the musical representation data - for example, from another system such as CALMUS. Here it further acts as a sound engine, which turns the musical representation into sound data via audio samples or different types of synthesis (e.g. additive, subtractive, granular, concatenative, or FM synthesis) or another sound generating technique. (See Appendix B.)

9.4. GAMEPLAY DEPENDENCY

The horizontal strategies for adaptive music in games differ from those of linear music. Where the horizontal flow of linear music can be composed independently from any extramusical interference other than the usual preconditions posed by the instruments, musicians, musical culture and so on, adaptive game music must react to the game in which it is set. This is a basic concern caused by the non-linearity of the computer game medium. The horizontal development of adaptive VDM in games can be either *dependent* or *independent* in relation to gameplay.

9.4.1. GAMEPLAY-DEPENDENT DEVELOPMENT

Gameplay-dependent generation of horizontal structure in this respect is generally bound to the horizontal development of the game - a non-linear development, but nonetheless a horizontal one - because game states and events occur and take over from one another sequentially in the course of time. Gameplay-dependent strategies thus generate horizontal development based on implementation, gameplay and player input.

Simultaneous occurrences of game states and game events must also be taken into account. This sort of game state or game event simultaneity is a very significant aspect in relation to adaptive VDM in games, as simultaneously occurring game states and events must be taken into account in the adaptive music implementation strategy. On one hand, this is necessary to avoid potential conflicts caused by several simultaneous data streams from RPCs, triggers and cues in the gameplay attempting to direct the compositional algorithm in several potentially conflicting directions at once. This suggests the need for a hierarchy of prioritization amongst data input streams. On the other hand, such game state or game event simultaneity can be utilized creatively to facilitate vertical musical structures through the careful mapping of said data streams to musical parameters. For instance, a game state called ”night time” might be mapped to generate slowly evolving legato clusters within a certain register range and of a certain harmonic structuring scheme, while at the same time another game state called ”low danger level” is active, which instructs the algorithmic compositional system to orchestrate these clusters as woodwind instruments within a certain velocity range. Simultaneously, location based triggers associated with an in-game forest may generate a narrow low register string cluster. As evident from this example, the orchestration of the adaptive music can be provided by the mapping of game data to musical parameters - effectively turning the mapping task into a task of adaptive orchestration.

These strategies must therefore take into account this dictation by the game by defining the way the music is concretely implemented in regard to triggers, cues, mapping, encoding and so on while considering how the game may be played by the player(s). Thus, in the task of adaptive VDM conceptualization and design, the concrete implementation as well as considerations on potential horizontal directions inherent in the offered gameplay have critical significance to the horizontal development of the generated music. This demands of the implementation task the inclusion of a strategy for horizontal development of the music.

9.4.2. GAMEPLAY-INDEPENDENT DEVELOPMENT

In addition, horizontal strategies may be conceptualized, which are *independent* in relation to gameplay in that they are not adaptively tied to game states or events.

Their functions in the game may include introducing musical variation and controlling the balance between presence or absence of music, and any adaptive coupling to the game they might have is limited to following very general aesthetic goals (such as maintaining vertical dominance or ensuring that the generated music represents the general pace and temperament of the game level.) As such, they represent a gameplay-independent aesthetic of the musical flow.

Gameplay-dependent and gameplay-independent methods for horizontal development will likely co-exist and work simultaneously. In such cases the dependent part of the system should have higher priority than the independent approach so that sudden, and perhaps important, events in the game take control over horizontal development when necessary. It is therefore crucial that the design of a horizontal strategy for VDM composition in games is concerned with both of these methods.

9.5. METHODOLOGY OF MAPPING HORIZONTAL VDM DEVELOPMENT

Horizontal change applied to vertical parameters (e.g. the evolution of pitch, instrumentation or amplitude over time) entails a mapping of some controlling agent to said parameters. As so far established, these agents of control may be seen as gameplay-dependent or gameplay-independent multidimensional envelopes. Possible methodologies for mapping such envelopes to vertical parameters may include *deterministic, indeterministic* and *stochastic* approaches.

9.5.1. DETERMINISTIC MAPPING

A deterministic approach here represents the most direct relationship between the controlling envelope and the parameter change. Here it may be expected that a certain value from the controlling agent (a particular point on the envelope's Y-axis) will consistently result in a particular value being generated for the vertical parameter every time. If, for example, a space pilot in a game like EVE Online approaches a wormhole and the distance between the spaceship and the wormhole is mapped deterministically to the amplitude and register range of a French horn chromatic cluster (so that amplitude and register range increase as the distance decreases), then one can expect the amplitude and register range of the cluster to be set identically every time the distance is at a certain value.

A deterministic mapping approach is accompanied by an inherent risk of creating a sense of repetition or even of pulse as repetitive game events create the same horizontal reaction every time they occur. Also, since the mapping represents a one-to-one relationship there is a potential for the underlying adaptive system to be revealed and the trans-diegetic function thereby to become *overt*. As discussed in chapter 6 this is of significance because overt trans-diegetic adaptive music (i.e. adaptive music which is implemented in such a way that the adaptive system itself becomes obvious) can enable the player to predict what the music will do when certain events occur in the game. And, more importantly perhaps, it can enable players to predict what will happen in the game based on the music. This can be highly problematic in a horror game, as mentioned before, where the music could accidentally give away the sudden occurrence of a monster before it happens - ruining the surprise of the event.

9.5.2. INDETERMINISTIC MAPPING

Indeterministic approaches may involve random generators that are set to output values within a specified range. In the prototype, KeynotePlaybackSystem (KPS), the choice of sound file to play within each of two available ”threat level” categories is indeterministic as it chooses randomly between 10 files. The envelope generators for [High Reg] and [Low Reg] (high and low register respectively) are likewise indeterministic in their choice of attack, hold, release and pause time values within a set range. Additionally, the KPS prototype's functionality for random pitch variations functions indeterministically within a pitch range that can be set in semitones. (See Appendix D)

Indeterministic approaches suit VDM well because the random processes tend to avoid the formation of *perceivable horizontal regularities*. VDM's aspect of *un-specificity* also goes well in hand with indeterministic approaches on the level of the musical output. Let me return to the adaptive coupling of register range in the recent French horn cluster example to an RPC that represents the distance to a wormhole. Here it may be more significant that the register range increases to within a certain range as the spaceship approaches than how much it precisely increases. Indeterministic mapping between controlling envelopes and vertical parameters has a high potential in this way for avoiding a sense of repetition and regularity. (Methods to avoid regularities are investigated practically in appendices A, B, C and D.) Additionally, this approach may help avoid the trans-diegetic function becoming *overt*.

9.5.3. STOCHASTIC MAPPING

Stochastically based temporal development entails the use of probability calculations from which to derive horizontal structure. While I do not have a prototype at the time of writing that offers functionality based on probability, one may easily imagine the great applicability of this method.

Probability-based control may allow RPCs (e.g. the time of day, damage taken or threat level), game events (e.g. the player's spaceship is hit by a missile) or direct player inputs (e.g. if the player clicks on a certain button) to introduce spectromorphologies, instrumental groups, pauses and so on in a non-patterned and irregular way. But opposed to indeterministic mapping, which can be made to choose at random within a specified range, stochastic mapping can increase or decrease the *likelihood* of some change to occur. If the spaceship approaches the wormhole and instead of the register range of the horn cluster increasing it is the probability of the onset of, for example, a myriad of short outbursts of deep contrabass tremolos to occur that increases, then vertically dominated musical intensity can be said to build up in a stochastically facilitated manner. As also applies for indeterministic mappings, it is more likely that trans-diegetic couplings can be kept covert through stochastic methods than is the case with deterministic mappings.

Very interestingly, being able to set the probability of the occurrence of horizontally separable gestalts, horizontal regularities (such as repetition or pulse) and degree of spectromorphological bonding would effectively enable, for example, an RPC to *control the degree of vertical dominance* of the output music. Such a feature would entail quite a bit of programming and require a hybrid system of generative and VE-transformational algorithms to be developed, but the game-music functional potential could very well be worth the effort. The potential of being able to control the degree of vertical dominance has the perceptual implication of allowing for control over the anticipatory listening position of the player through a single variable. This would be a highly attractive function for an adaptive music-generating algorithm in games. If, for example, there is a presence of *"something" that exceeds the boundaries of the senses* (for instance, a ghost) in some parts of a game's topology while not in others, then the position of the player in the game world could control the music's dominance by either the horizontal or vertical dimension in synchrony with such a presence.

9.6. CONCLUSION

Vertical dominance is encouraged by the three characteristics of VDM (to *avoid horizontally separable gestalts, perceivable horizontal regularities*, and *weaken the perceived bond between spectromorphologies*). When these characteristics are treated proactively as a basis for composition, it can be useful to know how they relate to horizontal development. The stringency with which these characteristics must be followed by the composed musical structure changes, I have argued, according to the time scale at which these structures reside. The relationship between time scale and the necessary stringency to comply with the three characteristics of VDM is characterized by the principle that the lower the time scale, the higher the stringency.

As such, structures at the *vertical extension* time scale must comply strictly to these characteristics. Because the auditory verticality of the vertical expression per definition dictates *stasis*, no perceivable development can take place at this time scale.

Structure at the *sound object* time scale must avoid structures that will cause the listener to form separate sound object gestalts. But as opposed to structures at the vertical extension time scale horizontal development is allowed within the sound object as long as it maintains the integrity of the coherent spectromorphological gestalt.

Meso time scale structures must tend towards perturbation-based anticipation by being unpredictable, highly entropic, containing no perceivable repetition or causal direction and by taking place through irreversible change. They may be constituted by both smoothly overlapping and clearly separable sound objects depending on the desired degree of vertical dominance.

Macro time scale structures may well have internal pauses and be constituted by clearly separable sound object- and meso-structures. However, also at this time scale, if the aim is to protect the integrity of vertical dominance, horizontal formations may not violate the rules associated with perturbation-based musical anticipation by exhibiting obvious patterns, predictability or repetition and by spectromorphological bonding being at a minimum.

At the *supra* time scale, which in games may be regarded to represent an entire game or game series, the characteristics of VDM are effectively insignificant due to the magnitude of the time scale and its detachment from the immediate gameplay and listening experience.

The musical representation associated with a specialized algorithm for computer-generating VDM in games may be conceived as *multi-dimensional meta-structural envelopes*. Envelopes in this respect are effective form-shaping tools, which can be deployed on vertical parameters and any accessible data stream (such as RPCs) can potentially be used as input to define the progression of such an envelope. As a means for achieving horizontal development envelopes may be *static* or *adaptive* by being fixed shapes or by being generated in real-time respectively.

A multidimensional envelope may serve as a musical representation better suited for computer-generated VDM than a score-based representation as it may allow for precise control over the kind of gradual and continuous movements associated with irreversible change.

I have identified in this chapter that in addition to functioning as a means for *composition* by controlling vertical parameters over time, envelopes may also be used to enrich the sounding musical manifestation by *superimposing* micro time scale irregularity onto material generated by for example a generative algorithm like CALMUS.

The horizontal development of adaptive VDM in games, I have argued, can be either *dependent* or *independent* in relation to gameplay. Gameplay-dependent development is tied to the gameplay. A range of game states and events may be happening at the same time and all of them may potentially have an influence on the gameplay-dependent music's evolution. A horizontal strategy must therefore contain a hierarchy of prioritization amongst data input streams. In such cases the mapping task can effectively be regarded as a task of adaptive orchestration. Gameplay-independent development entails that the evolution of the music happens by other means than the gameplay. Such developments are not tied to game events and game states but may represent general aesthetic goals such as maintaining vertical dominance or following the general pace and temperament of the game level. It is crucial that the design of a horizontal strategy for VDM composition in games is concerned with both gameplay-dependent and gameplay-independent developments as these may very likely come to co-exist.

The methodological possibilities for the mapping of envelopes (whether static or adaptive, gameplay-dependent or gameplay-independent) to musical parameters include deterministic, indeterministic and stochastic approaches.

A deterministic approach represents a one-to-one relationship between the controlling agent and the resulting parameter change. It holds the risk of causing a sense of repetition and pulse. Additionally, a deterministic approach is likely to reveal the adaptive system and make any trans-diegetic functionality *overt.* This can cause narrative problems since the player might then be able to predict game events based on the music - an inappropriate side effect in some cases.

Indeterministic approaches are based on random choices within specified ranges. Such an approach suits VDM well because it tends to avoid the formation of *perceivable horizontal regularities* and encourage *un-specificity.* It has the added benefit of encouraging a concealment of the underlying adaptive system. Thereby, trans-diegetic functionality is easier to keep covert.

Stochastic approaches to the mapping of in-game data to musical parameters have the same effect on trans-diegetic material. Here it is probability calculations that act as the basis for horizontal musical development enabling for instance RPCs or direct player inputs to control the likelihood of certain musical spectromorphologies or other musical events to occur.

As an ultimate potential, which would require a great deal of programming to realize, a stochastic approach might be used to control the probability of the occurrence of horizontally separable gestalts, horizontal regularities and degree of spectromorphological bonding. This would effectively allow a single data stream from a game to control the degree of vertical dominance in the generated music. This would be a highly attractive trans-diegetic function in games that have as part of their narrative such phenomena as outer space, infinity, mystery, fear, divinity, paranormal activity or generally a presence of “something” that exceeds the boundaries of the senses.

CHAPTER 10. CONCLUSION

In this thesis, I have explored the implications of implementing vertically dominated music in computer games. The focus has been on games that have as part of their narrative such phenomena as outer space, infinity, mystery, fear, divinity, paranormal activity - or generally, a presence of "something" that exceeds the boundaries of the human senses.

Through addressing theoretical as well as practical aspects of the subject, I have provided a structural, aesthetic and functional theoretical foundation from which an automated compositional system may be built. I have developed a phenomenological, semiotic and music-perceptual framework of understanding, and through this framework, assessed the narrative, semiotic and perceptual implications on the gaming experience of introducing VDM in a computer game. This has provided the theoretical basis necessary to formulate an initial set of vertically dominated compositional principles and put these principles to practice as a means for demonstrating the application of the developed concepts.

Chapter 2

Chapter 2 served as a theoretical substrate for many aspects of the thesis. Its contribution was focused on **enquiry 1 of the problem statement** by addressing the fundamental problem of *verticality*.

I argued in this chapter for the differentiation of Jonathan D. Kramer's concept of *vertical time* into *cyclic* and *non-cyclic vertical time* governed by *cyclic* and *non-cyclic musical stasis* respectively. Additionally, I found it necessary to append to these musical time conceptions the notion of *irreversibly changing vertical time*. Here temporal development clearly plays a part in the musical expression, but it does so through continuous movement and directed toward no predictable goal. I made the case that *auditory verticality* exists as a percept only as a coherent, *unstriated* and *unquantified* whole with no beginning and no end. In the experience of such a whole, the listener does not *segregate sequentially* the continuous flow of sound. I proposed to address the temporal delimitation of the perceptual phenomenon of *auditory verticality* as a temporal extension of the vertical - a *vertical extension.* Perceptually speaking, the contents of the *vertical extension*, I proposed, can be seen as an expression of Bergson's *pure duration* and Boulez' *smooth space-time.* Its constituting sound material is continuous rather than discrete and promotes an illusion of *auditory stasis.* I furthermore extended Curtis Roads' time scale model with an additional time scale - the *vertical extension time scale*, which encompasses the micro time scale but may likely exceed it.

I provided a definition of *auditory verticality* as:

> **A perceptual phenomenon, which is facilitated by an illusion of "auditory stasis" caused by structural properties of a "vertical**

> **expression” and occurring within a temporal window of a ”vertical extension”.**

I subsequently defined *Vertically Dominated Music (VDM)* as:

> **Music in which the expressivity of auditory verticality dominates over that of auditory horizontality.**

Auditory horizontality was in turn defined as:

> **A perceptual phenomenon of those auditory developments that exceed the ”vertical extension” and which facilitate the formation of clear temporally truncated gestalts.**

And Horizontally Dominated Music (HDM) as:

> **Music in which the expressivity of auditory horizontality dominates over that of auditory verticality.**

Additionally, I proposed in this chapter that what happens, at the perceptual level, within the duration of the vertical extension can be regarded as an experienced fusion of temporal and spatial properties into an *auditory singularity* in which horizontal and vertical structures are not perceptually separate.

Chapter 3

Chapter 3 addressed subjects that relate to primarily **enquiries 1 and 2 of the problem statement** by looking into the structure of auditory verticality - that is, the structure that governs the vertical extension.

As a contribution to the structural part of the theoretical foundation put forth in the thesis, I provided a definition of the notion of the *vertical expression* understood as: *an emergent hybrid concept of pitch, harmony, timbre and micro-texture*.

As a necessary part of the development of an initial set of compositional principles that are aimed at VDM, I identified a list of *vertical parameters* that may be used for vertically dominated compositional purposes in a game context. These parameters are listed in appendix F and were categorized into five main aspects: *tone systems, harmony, timbre, micro-texture* and *aural space* - of which aural space was discussed more peripherally.

Chapter 4

Chapter 4 was likewise concerned primarily with **enquiries 1 and 2 of the problem statement**. The chapter contributed further to the theoretical foundation by addressing structural and aesthetic aspects of VDM. It provided insight into some techniques of VDM composition that may be used in an initial set of compositional principles, and expanded the list of vertical parameters initiated in chapter 3.

I argued that VDM, as an umbrella term, can be associated with music across styles and historical periods on the basis of common underlying structural and perceptual properties.

I provided a crucial basis for formulating an initial set of compositional principles by identifying three general characteristics that facilitate vertical dominance:

- **Avoidance of horizontally separable gestalts**
  - By obscuring of onset and termination
  - By mass-merging of many individual events
- **Avoidance of perceivable horizontal regularities**
  - Due to an absence of repetition, rhythmic pulse and accentuation of meter
  - Due to an absence of patterns like themes, motifs and rhythmic patterns
- **Weakening of the bond between spectromorphologies**
  - By an absence of causal direction
  - By exceeding the period of event cohesion

A set of techniques for the expansion of vertical expressivity was listed that includes:

- New orchestral playing techniques
- Emphasis on mass-effect rather than individual pitches and rhythms
- New musical texture conceptions
- Experiments in timbre
- New orchestration techniques
- New technologies
- New horizontal development schemes
- Focus on the perceptually *unspecific*

I defined *musical entropy* on the levels of harmony, melody, timbre and rhythm. The degree of musical entropy was tied to the degree of vertical and horizontal dominance. I applied the notion of the *rhizome* as a metaphor for VDM structure due to its *non-hierarchical, a-centered* and *non-linear* characteristics that resemble those of auditory verticality (as opposed to arborescent, tree-like structures, which represent hierarchy, causality and an underlying generative model.)

A look at spectral music offered a range of ideas and music compositional approaches - some of which can inspire structuring schemes for adaptive VDM composition. This led me to add *ring-modulation, spectral stretching and compression* and *spectral pitch-shifting* to the list of vertical parameters presented in Appendix F. I introduced the idea that the concept of *reservoirs,* used by spectral composers, may be utilized to create textural templates that can be superimposed or used to co-create vertical expressions in an initial set of compositional principles.

I argued that the inaccuracy of intonation in orchestral synthesis is at the same time a challenge and a defining factor of the style, and that in the context of entirely computer-based VDM, orchestral synthesis holds the potential to become pitch-perfect without leaving behind the orchestral instrumentation. I identified a tendency for perceptual *un-specificity* as a common denominator in the VDM styles presented in the chapter. On one hand, the included composers have an impressive detail in the written score, while, on the other hand, they use this detail to accomplish textures, in which those details are in and of themselves largely insignificant to the overall musical expression.

A significant observation in regard to formulating an initial set of compositional principles was also that the need for *specificity* is primarily tied to the presence of an orchestra. If one takes away the orchestra, this need disappears with it.

Chapter 5

The subject of chapter 5 was primarily concerned with **enquiry 3 of the problem statement**. The chapter took important steps towards presenting a phenomenological and music-psychological framework of understanding, which, together with an examination of VDM's narrative potential in chapter 6, formed the basis for determining the impact of VDM on the perception of a game or other interactive experience. It did so by exploring the perception of VDM, and, not least, by devising a model for understanding music listening in the context of VDM as well as more broadly.

In this chapter, I proposed a theoretical model for explaining aspects of music perception that have been highly significant throughput the thesis. I referred to this model as "*three anticipatory listening positions*". It comprised *equilibrium-based anticipation* (or future-listening), *historically based anticipation* (or past-listening) and *perturbation-based anticipation* (or now-listening).

I suggested that the perception of VDM gravitates toward the extreme position of *perturbation-based anticipation*, by tending to cause a constant cognitive perturbation and by forcing the listener to constantly adapt his or her schemas of the music through accommodation rather than assimilation. This position in the model ideally requires that the listener is unfamiliar with the music, and that the music is very unpredictable by exhibiting a high degree of musical entropy, no causal direction, an absence of repetition, and, finally, by vertical dominance.

I argued that the perceptual attributes of VDM include a potential feeling of *control loss* connected to VDM's special quality of moving in and out of *musical listening* - thus playing with the listener's ability to *empathetically co-enact* the musical structures as they occur.

I took the position that music regarded as ”a state of listening” holds strong potentials for shedding light on the perceptual workings of VDM. From this perspective, I proposed regarding the notion of ”music” as: *That* which is experienced when in a state of *musical listening*. I proposed to understand *musical listening* as *empathetic co-agency,* which describes a participatory action in which the listener, through empathetic capabilities tied to the mirror neuron system, enacts the music while listening.

I introduced the term *empathetic co-agency* as combining linguistic, motor-functional, musical and emotional aspects of mimicry allowed by the human mirror neuron system. And I argued that *empathetic co-agency* is partially dependent on various degrees of hierarchical structures, patterns and musical horizontality or other means with clear musical associations. I could subsequently make the point that both on a music-cognitive level as well as at the level of embodied experience, a direction in the model towards perturbation-based anticipation is associated with a sense of control loss, whereas the direction towards equilibrium-based anticipation lets the listener gain a sense of control over the empathetic co-action.

I addressed the question of an undisclosed, unidentified, non-concrete and perhaps illusory *sender*, who takes the role as an origin of *intentions* within a linear sender-medium-receiver communicational paradigm. The notion of a *sender* in this respect was seen as occupied with a certain *state* constituted by *intentions* that are communicated through music and with which the receiver co-enacts empathetically. This empathetic co-agency with a *sender* is, I suggested, intrinsically present in the *musical listening* state. This led to a crucial question in regards to VDM perception:

> **What is the state of an undisclosed sender, whose motor-functional, linguistic and emotional properties, when mediated through sound, produce VDM?**

I proposed that the perception of VDM may be partially understood as the perceiver’s empathetically co-enacted ”answer” to this question - in the form of a mental representation based on mimicking capabilities of the mirror neuron system that facilitates a realization of the sender’s state through music listening. I further argued that this role of a ”sender of intentions” may be attributed by the perceiver to narrative elements of a multi-modal percept.

Based on the findings of the chapter, I was able to condense the perceptual attributes of VDM into the adjectives: *continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless* and *unpredictable*. I argued that emotional attributes of VDM may be of positive or negative valence and that it is the *vertical expression* that is primarily responsible for the polarity of this valence. The perceptual attributes of vertical dominance here play a highly significant secondary role, adding their own unique qualities that are associated with the above adjectives.

Chapter 6

In chapter 6, I looked into the narrative potential of VDM. This chapter covered subjects that have contributed to answering **enquiries 1, 3 and 4 of the problem statement**. Clarifying VDM's narrative potential was a key component in the effort to create a structural, aesthetic and functional foundation from which a generative VDM system may be built. I approached this by addressing issues related to diegetic layers, semiotic and perceptual attributes of VDM, and multi-modal perception - all of which have also contributed to the understanding of the impact it would have on the gaming experience to implement VDM. The chapter further provided an exploration into possible deployments of its theoretical conceptions by proposing cross-modal coupling possibilities for VDM in games.

It is problematic to speak about a non-diegetic narrative layer in games because games are interactive. I suggested, however, that players are likely to project the presence of such a layer from the film media, and argued that the notion of non-diegetic music is therefore highly useful in "film-like" games. The inherent mystery of the non-diegetic layer, what I referred to as "the great beyond" of the non-diegetic, further supports, I proposed, the "undisclosedness" of non-diegetic music's *undisclosed sender* in cases where a linear sender-music-receiver communicational paradigm is phenomenologically maintained because any trans-diegetic functionality is *covert*. The narrative potential of non-diegetic music, may on the other hand be weakened in cases where a non-linear, circular communicational paradigm exists, due to phenomenologically *overt* trans-diegetic functionalities.

I argued that VDM, and music in general, may be seen to narrate through what I called *competence-related semiotic attributes* and *empathetically co-enacted perceptual attributes* that are projected onto elements of the multi-modal construct that the music couples to. In the case of VDM, this projection includes this music's established association to horror and mystery in the competence of audiences and players - as well as its perceptual attributes, which I condensed to the adjectives *continuous, limitless, timeless, chaotic, infinite, incomprehensible, uncontrollable, unknown, weightless, homeless* and *unpredictable*.

I proposed that couplings may be made between music and game in a *general* or a *specific* manner. A *general* coupling implies that the mentioned semiotic and perceptual attributes of VDM are projected onto the *overall* narrative and aesthetics of any multi-modal phenomenon. *Specific* music-structural constituents of the style (such as certain instruments, playing techniques etc.) may, on the other hand, also be coupled to *specific* in-game elements (such as characters and locations.) I further suggested the terms *game-internal* and *game-external* competence to denote veridical and schematic competence-related expectations respectively.

I identified, in the chapter, that a VDM based game-musical concept has to build on different musical structuring schemes than HDM. I argued that such vertically oriented structuring schemes may serve as vertically dominated equivalents to, for example, the horizontally dominated notion of the leitmotif, and do so based on the vertical parameters identified in the thesis.

In this chapter, I also looked at VDM's relation to space-based science fiction. Here I proposed that VDM holds in common with, for example, the use of theremin in the 40's and 50's a tendency for breaking with the traditional musical striation paradigms as a means to convey both altered states of consciousness as well as outer space as a narrative setting. I highlighted that VDM's association with scenarios of horror is characterized by dissonant and negatively valenced vertical expressions, while what I referred to as *benign* VDM can be seen to convey meditative peace and reflection.

Infinite reverb was discussed. I regarded it as an expression of unstriated ”space” rather than striated ”room” that shares some perceptual attributes with VDM such as infinity, incomprehensibility, continuity and limitlessness through a breaking with logic and corporal limits of embodied experience.

I further made the point that VDM, as experienced through empathetic co-agency, can be seen to break with the boundaries of an *inner space* in order to convey *outer space.* And that this represents a special narrative ability of VDM that is of great significance in the context of this thesis. I proposed that breaking with the culturally based and top-down processed framework of our musical competence through vertical dominance, also entails an ability to break with the cognitively fundamental and natural bottom-up processed preconditions of human perception.

As a welcome response to my initial fascination with the music in Kubrick's films mentioned in the introductory chapter, I could identify that it is not least *this* special ability that lets VDM so effectively give a voice to such phenomena as outer space and a presence of ”something” that exceeds the boundaries of the senses.

Chapter 7

Chapter 7 addressed general implication of sound and music in computer games. It stands as a significant pillar in the theoretical foundation for an automated compositional system of VDM to rest upon by examining, not least, the functional aspects of music implementation in games. Additionally, I identified in this chapter a range of issues that an initial set of compositional principles must adhere to when implemented adaptively in games. Chapter 7 thus contributed primarily to **enquiries 1 and 2 of the problem statement**.

I highlighted that linear music implemented in a non-linear medium will fail to perform most game-musical functions. However, non-linearity is scalable and some game genres are more non-linear than others. I made the case that sandbox box games and MMOs will often have a particularly high degree of non-linearity and that the compositional challenges posed by non-linearity are therefore particularly strong in such games. To synchronize to an indeterminate course of events, and thereby give access to a number of game-musical functions, game music must be adaptive. It can be so in either a *direct* or an *indirect* fashion by reacting to direct player inputs or indirect game states respectively (corresponding to what Karen Collins has called interactive and adaptive dynamic music). Three primary pre-requisites were identified for adaptive music in games: Flexibility, variation and trans-diegetic coupling. The avoidance of listening fatigue as well as the definition of triggers, cues and RPCs to couple the music to were also highlighted as central concerns.

Branching, layering, transitions and generative music are common approaches to achieving adaptive music. I proposed that layering is a suitable technique for VDM because it is basically a vertical method; that transition modules are to a large extent unnecessary in VDM because the problems they address are of a horizontally dominated character. Branching, being an expression of an *arborescent* design structure, entails hierarchy, linearity and limitations in regards to possible transitions. In opposition to the tree-structure, I presented the Deleuzo-Guattarian conception of the *rhizome* in which all points ideally connect to all other points. I regarded the rhizome structure as an ultimate expression of non-linearity that may represent some game designs such as persistent universe, sandbox type games that have a high degree of non-linearity. I further proposed that music with a rhizomatic structure, such as VDM, is therefore structurally well suited for being implemented adaptively in such games.

Achieving variation was identified as a central concern in adaptive computer game music. An adaptation of Karen Collins' list of variation parameters was presented in the context of adaptive VDM. These were seen as directly applicable as part of an initial set of compositional principles for adaptive VDM in games.

Immersion and *presence*, which are generally regarded as key goals in game design, may be counteracted by linear music because it may present an alternative to the game's virtual reality rather than supporting it. Additionally, unprovoked abrupt transitions between musical cues stand for an increase in critical distance and thus a weakening of immersion. Presence is likewise threatened by anything that will bring attention away from the involvement in the virtual reality and onto the media as a media. I suggested that the level of listening attention referred to by Truax as *background listening* is less likely to break with immersion than those of *listening-in-search* and *listening-in-readiness*.

I synthesized a non-comprehensive list of seventeen functions that may be performed by VDM. The list was based on musical functions in game music, sound design of an acoustic ecology and film music and stands as an important aspect of the theoretical foundation for the development of an automated system for the implementation of adaptive VDM in computer games.

I concluded this chapter by highlighting that the unique structural and perceptual attributes of VDM places it in an advantageous position as compared to HDM in regards to all of the above-mentioned game-related implications. I also pointed out, however, the existence of some technological challenges that are not least associated with VDM's tendency for very large-scale polyphony and demands for a high degree of complexity in its *vertical expression*.

Chapter 8

Chapter 8 was devoted to establishing a functional and aesthetic contribution to the theoretical foundation for adaptive VDM in games by addressing both advantages and challenges that arise from realizing music via computers through algorithmic techniques. It further contributed to paving the way for an initial set of composition principles by providing an assessment of the best algorithmic approach for the task. The chapter is thus concerned mostly with **enquiries 1 and 2 of the problem statement**.

I proposed that there is a direct relationship between what I termed the *basic unit of composition,* with which a system operates, and the degree of potential for the narrative interconnectedness between the gameplay and a game's music. While a distinction between *transformational* and *generative algorithms* proposed by Wooller et al. is useful, I argued that from a structural perspective this categorization scheme is problematic. I therefore suggested a different scheme, which is based on the time scale at which the organization of musical data is taking place - a system's basic unit of composition. This resulted in a scale ranging from linear music over transformational and generative algorithms to *VE-transformational* and *micro-generative algorithms.* I introduced the two latter terms to account for a more radical approach to algorithmically composing VDM.

I highlighted that generative approaches to adaptive music in games entails serious aesthetic challenges that inhibit its implementation in AAA games. I focused on two significant aesthetic challenges of generative music in this respect; *musical coherency* and the *complexity of the sounding musical manifestation*. Additionally, I could identify that by not offering access to the vertical extension or micro time scales the traditional techniques of generative music are not to any serious extent capable of generating the most crucial aspect of VDM, the *vertical expression*. Furthermore, the rhizome-like structure of VDM speaks against a generative method as rhizomes per definition are not amenable to a generative model.

The *music-shaping stages* differ greatly between the production pipeline of traditional acoustic music and computer-generated music. The nature of computer-generated music is such that its music-shaping stages have to all be designed. This means that computer-generated music with a high degree of control lacks the "free" additions of musical complexity offered by, for instancem acoustic orchestral music. For VDM, in which the complexity of the sounding now - the vertical expression - is of utmost importance, this challenge, I argued, is particularly serious. I proposed that a specialized system for adaptive VDM must therefore have the *vertical expression* as its lowest unit of organization in order to have a fitting musical predisposition. I further elaborated that VDM thus demands the most adaptive and therefore also most aesthetically challenging methods in terms of musical coherency and complexity of the sounding musical manifestation: namely *VE-transformational* and *micro-generative* algorithms. A VE-transformational approach is not capable of facilitating certain continuous movements such as glissandi due to its ”modular” design scheme of organizing static vertical expressions.

However, I concluded that a VE-transformational approach with limited micro-generative capabilities, which will enable it to compose an irreversibly changing musical flow through most vertical parameters, is well suited both for game music and as a specialized system for composing VDM.

Additionally, I mentioned that transformational and generative techniques, while not being ideal may be expanded to perform better as basis for VDM by introducing control over the vertical extension and micro time scales.

Chapter 9

Chapter 9 contributed to the theoretical foundation of the thesis by exploring possibilities for horizontal developments in adaptive VDM. Looking into the temporal aspect of VDM composition represents a vital aspect in the construction of an initial set of compositional principles for adaptive VDM. The chapter also addressed practical aspects of narrative couplings between music and game. As such, the focus was here on **enquiries 1, 2 and 3 of the problem statement**.

I looked at the relationship between time scales and the necessary stringency to comply with the three characteristics of VDM. This relationship is characterized by the principle that *the lower the time scale, the higher the stringency*. In this way, structures at the *vertical extension* time scale must comply strictly to these characteristics while they are effectively insignificant at the *supra* time scale.

I suggested that the musical representation associated with a specialized algorithm for computer-generating VDM in games may be conceived as *multi-dimensional meta-structural envelopes.* And that as a means for achieving horizontal development, envelopes may be *static* or *adaptive* by being either fixed shapes or generated in real-time respectively. I proposed that a multidimensional envelope may serve as a musical representation better suited for computer-generated VDM than a score-based representation as it may allow for precise control over the kind of gradual and continuous movements associated with *irreversible change.* In addition to functioning as a means for *composition* by controlling vertical parameters over time, envelopes may also be used to enrich the sounding musical manifestation by *superimposing* micro time scale irregularity onto material generated by, for example, a generative algorithm like CALMUS.

The horizontal development of adaptive VDM in games, I argued, can be either *dependent* or *independent* in relation to gameplay. A horizontal strategy must therefore contain a hierarchy of prioritization amongst data input streams, and the mapping task can effectively be regarded as a task of adaptive orchestration. Furthermore, it is crucial in this respect, that the design of a horizontal strategy for VDM composition in games is concerned with both gameplay-dependent and gameplay-independent developments as the two may very likely co-exist during a game.

It was highlighted that the methodological possibilities for the mapping of envelopes to musical parameters include *deterministic*, *indeterministic* and *stochastic* approaches. A deterministic approach represents a one-to-one relationship between the controlling agent and the resulting parameter change, and holds the risk, I argued, of causing a sense of repetition and pulse as well as revealing the adaptive system and thereby making any trans-diegetic functionality *overt.* This can cause narrative problems since the player might then be able to predict game events based on the music - which may or may not be desirable. Indeterministic approaches, which are based on random choices within specified ranges, suit VDM well because they tend to avoid the formation of *perceivable horizontal regularities* as well as encourage *un-specificity.* I argued that an indeterministic approach has the added benefit of concealing the underlying adaptive system, whereby trans-diegetic functionality can be kept covert. Stochastic approaches to the mapping of in-game data to musical parameters have the same effect on trans-diegetic material. Here it is probability calculations that act as the basis for horizontal musical development

The chapter concluded by pointing to possible future work, by proposing the development of a stochastic system that can control the probability of the occurrence of horizontally separable gestalts, horizontal regularities and degree of spectromorphological bonding. This would effectively allow a single data stream from inside a game to control the degree of vertical dominance in the generated music - a highly attractive trans-diegetic function in games that have as part of their narrative such phenomena as outer space, infinity, mystery, fear, divinity, paranormal activity or generally a presence of “something” that exceeds the boundaries of the senses.

Appendices

A practical perspective was documented in the appendices A, B, C, D and E as a means for addressing **enquiries 2 and 4 in the problem statement**. These projects included: The prototype, NowEngine; the development of an irregular sound library for enriching the sounding complexity of a generative system; the composition of four linear VDM pieces titled DistanceMusic; the prototype, KPS; as well as experiments done with implementing VDM in EVE Online through the middleware program, Wwise. This work specifically concerned the deployment of research results in practice and, vice versa, functioned as inspiration to ideas and concepts developed in the thesis.

Initial Set of Compositional Principles for Adaptive VDM

Based on the research, I have identified the following primary aspects as constitutive for an initial set of compositional principles. These principles are tightly interconnected with the theoretical foundation presented in the thesis, and the summary below represents a highly condensed, structurally and practically oriented selection of the wider range of research results.

Firstly, in order to compose VDM an automated system must to various extents adhere to the general characteristics that facilitate vertical dominance by:

- Avoiding horizontally separable gestalts
- Avoiding perceivable horizontal regularities
- Weakening the bond between spectromorphologies

It can do so by: Obscuring the onset and termination of tones and spectromorphologies; mass-merging many individual events; avoiding repetition, rhythmic pulse and accentuation of meter; avoiding patterns like themes, motifs and rhythmic patterns; avoiding causal direction; and by exceeding the period of event cohesion.

It may utilize multidimensional envelopes to serve as its musical representation and use these envelopes to achieve horizontal musical development through any of the wide range of vertical parameters presented in Appendix F within the categories of tone systems, harmony, timbre, micro-texture and aural space.

Its manifestation may be carried out through synthesis or based on audio samples. In both cases it should exhibit micro-textural irregularities to ensure a sufficient complexity of the sounding musical manifestation. these irregularities should be implemented on a sound-engine basis rather than on a per sample basis.

A specialized system for adaptive VDM must have the *vertical expression* as its lowest unit of organization, but it must also be capable of facilitating irreversible change through, for example, glissandi. This makes a VE-transformational approach with limited micro-generative capabilities well suited as a specialized system for composing VDM in computer games.

Comment on the theoretical approach

In many cases I have used theoretical concepts and ideas in the research, whose theoretical context I have not necessarily adhered to - either because they were irrelevant to my focus or because some aspects of the theories have been problematized since their conception. This includes, for instance, the theory of Piaget's cognitive schema. I have taken the liberty, in such cases, to extract the useful aspects of such theories and put them to use autonomously.

Furthermore, I have found it necessary to apply a significantly broad range of perspectives to cover the subject matter of the thesis. This has the side effect that some aspects of the research are dealt with merely peripherally. Such matters include the notions of *immersion* and *presence* in games, the notion of *aural space* and some perspectives in the field of, for example, semiotics that would have been interesting to go further into. I have, however, sought to present a multifaceted approach in my research, and it has not been possible to treat all areas with equal emphasis.

While I cannot claim to having presented a comprehensive study on the subject, I have presented, in this thesis, an initial step toward understanding the implications of introducing VDM in a game context - a subject that, until now, has been largely unexplored, both in academia as well as in the computer game industry.

LITERATURE LIST

2K_Games. (2004). BioShock: 2K Games.

Abrams, J. J. (Writer). (2015). Star Wars: The Force Awakens: Lucasfilm, Bad Robot, Truenorth Productions.

Anderson, P. W. S. (Writer). (1997). Event Horizon: Golar Productions.

Assembly, C. (2014). Alien: Isolation: SEGA; 20th Century Fox.

Atmanspacher, H., & Filk, T. (2010). A proposed test of temporal nonlocality in bistable perception. *Journal of Mathematical Psychology, 54*(3), 314-321.

Atmanspacher, H., Filk, T., & Römer, H. (2004). Quantum Zeno features of bistable perception. *Biological cybernetics, 90*(1), 33-40.

Bartkowiak, M. J. (2010). *Sounds of the Future: Essays on Music in Science Fiction Film*: McFarland.

Bay, M. (Writer). (1998). Armageddon: Touchstone Pictures.

Bergson, H. (1910). Time and Free Will. 1889. *Trans. FL Pogson. New York: Macmillan*.

Bergson, H., & Pogson, F. L. (2001). *Time and free will: An essay on the immediate data of consciousness*: Courier Corporation.

Berndt, A., & Hartmann, K. (2007). Strategies for narrative and adaptive game scoring. *Audio Mostly*, 141-147.

Bernstein, L. (1976). *The unanswered question: Six talks at Harvard*: Harvard University Press.

Bharucha, J. J. (1994). *Musical perceptions* (R. S. Aiello, John A. Ed.): New York, NY, US: Oxford University Press.

Bharucha, J. J., Curtis, M., & Paroo, K. (2006). Varieties of musical experience. *Cognition, 100*(1), 131-172. doi:http://dx.doi.org/10.1016/j.cognition.2005.11.008

Birx, H. J. (2009). *Encyclopedia of time: science, philosophy, theology, & culture* (Vol. 1): Sage Publications.

Blesser, B., & Salter, L.-R. (2009). *Spaces speak, are you listening?: experiencing aural architecture*: MIT press.

Bois, M. (1980). *Iannis Xenakis : the man and his music ; a conversation with the composer and a description of his works / [by Mario Bois]* (Repr. ed.): Westport, Conn. : Greenwood Press.

Bordwell, D., Thompson, K., & Ashton, J. (1997). *Film art: An introduction* (7th 2003 ed. Vol. 7): McGraw-Hill New York.

Boyle, D. (Writer). (2007). Sunshine: DNA Films.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*: MIT press.

Bregman, A. S. (2005). Auditory Scene Analysis and the Role of Phenomenology in Experimental Psychology. *Canadian Psychology/Psychologie canadienne, 46*(1), 32.

Brolinson, P.-E., & Larsen, H. (1981). Rock. Aspekter på industri, elektronik och sound. *Stockholm. BrolinsonRock. Aspekter på industri, elektronik och sound1981.*

Brown, E., & Cairns, P. (2004). *A grounded investigation of game immersion.* Paper presented at the CHI'04 extended abstracts on Human factors in computing systems.

Bungie. (2001). Halo: Microsoft Studios.

Calvert, G., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes*: MIT press.

Cameron, J. (Writer). (1986). Aliens: Twentieth Century Fox.

Campbell, E. (2010). *Boulez, music and philosophy* (Vol. 27): Cambridge university press.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the national Academy of Sciences, 100*(9), 5497-5502.

Carroll, S. (2010). *From eternity to here: the quest for the ultimate theory of time*: Penguin.

CCP (Producer). (September 15th 2015). EVE Online Career Guide. Retrieved from http://go-dl.eve-files.com/media/1003/EVECAREERSGUIDE.pdf

Chion, M., Gorbman, C., & Murch, W. (1994). *Audio-vision: sound on screen*: Columbia University Press.

Christian Henson, J. H. a., & Smith, A. (2015) *Unpublished interview/Interviewer: J. Broomhall*. Game Music Connect 2015.

Collins, K. (2008). *Game sound: an introduction to the history, theory, and practice of video game music and sound design*: Mit Press.

Collins, K. (2009). An introduction to procedural music in video games. *Contemporary Music Review, 28*(1), 5-15.

Collins, K., Kapralos, B., & Tessler, H. (2014). *The Oxford Handbook of Interactive Audio*: Oxford University Press.

Collins, M. K. (2013). *From Pac-Man to pop music: interactive audio in games and new media*: Ashgate Publishing, Ltd.

Cope, D. (1993). *New directions in music* (6th ed.): McGraw-Hill Australia, Roseville, NSW.

Cuarón, A. (Writer). (2006). Children of men: Universal Pictures.

Cuarón, A. (Writer). (2013). Gravity: Warner Bros.

Deleuze, G., Guattari, F., & Massumi, B. (1987). A thousand plateaus: Capitalism and schizophrenia: University of Minnesota Press Minneapolis. (Reprinted from: Eleventh printing 2005).

Deutsch, D. (2013). *Psychology of music*: Elsevier.

Eigen, M. (1973). *Ingo Rechenberg Evolutionsstrategie Optimierung technischer Systeme nach Prinzipien der biologishen Evolution*: mit einem Nachwort von Manfred Eigen, Friedrich Frommann Verlag, Struttgart-Bad Cannstatt.

Ekman, I. (2014). A Cognitive Approach to the Emotional Function of Game Sound. In C. Karen, K. Bill, & T. Holly (Eds.), *The Oxford Handbook of Interactive Audio*: 'Oxford University Press'.

Eno, B. (1996). Generative Music. Retrieved from http://www.inmotionmagazine.com/eno1.html

Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. *Worlds in play: International perspectives on digital games research, 37*, 2.

ESA. (2015). Essential fact about the computer and video game industry. Retrieved from http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Prozesse. III. *Pflugers Arch Gesammte Physiol Menschen Thiere, 11*, 402-412.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17*(8), 1703-1714.

Fineberg, J. (2000a). Guide to the basic concepts and techniques of spectral music. *Contemporary Music Review, 19*(2), 81-113.

Fineberg, J. (2000b). Spectral music. *Contemporary Music Review, 19*(2), 1-5.

Fraisse, P. (1984). Perception and estimation of time. *Annual review of psychology, 35*(1), 1-37.

Friedkin, W. (Writer). (1973). The Exorcist: Warner Bros.

Gabor, D. (1947). Acoustical quanta and the theory of hearing. *Nature, 159*(4044), 591-594.

Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social neuroscience, 3*(3-4), 317-333.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593-609.

Gaver, W. W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological psychology, 5*(1), 1-29.

Gelder, B. d. (2006). Towards the neurobiology of emotional body language. *Nat Rev Neurosci, 7*(3), 242-249. Retrieved from http://dx.doi.org/10.1038/nrn1872

Gell, A. (1992). The anthropology of time: Oxford: Berg.

Gentry , J. (Writer). (2015). Synchronicity: POP Films.

Glass, A. L., Holyoak, K. J., & Santa, J. L. (1990). *Cognition*: Addison-Wesley.

Glassner, A. (2004). Interactive storytelling: Techniques for 21st century fiction.

Granat, Z. (2009). Rediscovering'Sonoristics': A Groundbreaking Theory from the Margins of Musicology. *Music's Intellectual History*, 821-833.

Grau, O. (2003). *Virtual Art: from illusion to immersion*: MIT press.

Gravesen, F., & Knakkergaard, M. (2003). Gads musikleksikon. In F. Gravesen, & Knakkergaard, M. (Ed.), *Gads musikleksikon* (1 ed.): Gads Forlag.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human brain mapping, 12*(1), 1-19.

Grimshaw, M. (2010). *Game Sound Technology and Player Interaction: Concepts and Developments: Concepts and Developments*: IGI Global.

Grimshaw, M., & Schott, G. (2007). Situating gaming as a sonic experience: the acoustic ecology of first person shooters.

Hannigan, J. (2015) *Unpublished Q&A session/Interviewer: K. R. Kristensen*. Game Music Connect 2015.

Harmonix. (2005). Guitar Hero: RedOctane.

Harmonix. (2007). Rockband: Electronic arts.

Hauge, L., & Thomsen, U. L. (1999). *Den nye psykologihåndbog*: Gyldendal uddannelse.

Hawking, S. (1988). *A brief history of time* (Bantam edition reissued 2011 ed.): Bantam Press.

Helmholtz, H. V. (1877). On the sensations of tone (English translation AJ Ellis, 1885, 1954): New York: Dover.

Hitchcock, A. (Writer). (1945). Spellbound: Selznick International Pictures, Vanguard Films.

Hobson, H. (Writer). (2015). Maggie: Lionsgate, Grindstone Entertainment.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*: U Michigan Press.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*: MIT press.

Husserl, E. (1966 [1887]). *The Phenomenology of Internal Time Consciousness*: Bloomington. INA: Midland Books.

ign. (2016). Eve Online. Retrieved from http://www.ign.com/games/eve-online-tyrannis/pc-62355

Iverson, J. (2010). The Emergence of Timbre: Ligeti's Synthesis of Electronic and Acoustic Music in Atmosphères. *twentieth-century music, 7*(01), 61-89.

Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and brain sciences, 31*(05), 559-575.

Jørgensen, K. (2008). Lyd som grensesnitt: Når dataspillets lyd bliver funksjonell. *MedieKultur. Journal of media and communication research, 22*(40), 9.

Jørgensen, K. (2010). Time for new terminology. *Diegetic and nondiegetic sounds in computer games revisited. Teoksessa Grimshaw, Mark (toim.): Game Sound Technology and Player Interaction: Concepts and Developments. Hershey, PA: Information Science Reference*, 78-97.

Karlin, F., & Wright, R. (2013). *On the track: A guide to contemporary film scoring*: Routledge.

Kassabian, A. (2001). *Hearing film: Tracking identifications in contemporary Hollywood film music*: Psychology Press.

Kirman, J. H. (1974). Tactile apparent movement: the effects of number of stimulators. *Journal of experimental psychology, 103*(6), 1175.

Knakkergaard, M. (2007). KLINGENDE TIDSRUM Om vilkår for opnåelse af indsigt i forhold mellem musik og tid. *Psyke & Logos, 28*(1), 22.

Knakkergaard, M. G., Finn. (2003). Gads Musikleksikon. In M. G. Knakkergaard, Finn (Ed.), (1 ed.): Gads Forlag.

Kokoras, P. A. (2005). *Towards a holophonic musical texture.* Paper presented at the International Computer Music Conference (ICMC).

Kramer, J. D. (1988). The Time of Music New Meanings, New Temporalities, New Listening Strategies.

Kristensen, K. S. R. (2011). A Musical Wormhole. 98. Retrieved from http://projekter.aau.dk/projekter/files/52680646/Specialet_Final.pdf

Kubrick, S. (Writer). (1968). 2001: A Space Odyssey [DVD]: Metro-Goldwyn-Mayer (MGM).

Kubrick , S. (Writer). (1980). The Shining: Warner Bros.

Lackner, J. R., & Teuber, H.-L. (1973). Alterations in auditory fusion thresholds after cerebral injury in man. *Neuropsychologia, 11*(4), 409-415.

Langkjær, B. (2000). *Den lyttende tilskuer: perception af lyd og musik i film* (Vol. 9): Museum Tusculanum Press.

Larson, R. D. (1985). *Musique fantastique: A survey of film music in the fantastic cinema*: Scarecrow Pr.

Leisen, M. (Writer). (1944). Lady in the Dark: Paramount Pictures.

Lerdahl, F., & Jackendoff, R. (1983). A Generative Theory of Tonal Music. 1983.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243*(4890), 489-494.

Ligeti, G. (1966). Über Form in der neuen Musik. *Darmstädter Beiträge zur neuen Musik, 10*, 23-35.

Ligeti, G. r. (1961). Atmosphêres. Universal Edition.

Ligeti, G. r. (1966). Lux Aeterna. Litolff/Peters.

Lissa, Z., & Tagg, P. (2015). Functions of Film Music. *Accessed February, 18*. Retrieved from http://www.tagg.org/teaching/mmi/filmfunx.html

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2), 0-0.

Lotze, M., Wittmann, M., von Steinbüchel, N., Pöppel, E., & Roenneberg, T. (1999). Daily rhythm of temporal resolution in the auditory system. *Cortex, 35*(1), 89-100.

Lucas, G. (Writer). (1977). Star Wars: Episode IV - A New Hope: Lucasfilm, Twentieth Century Fox.

Lynch, D. (Writer). (2006). Inland Empire: StudioCanal.

Mehra, J., & Rechenberg, H. (1999). Planck's Half-Quanta: A History of the Concept of Zero-Point Energy. *Foundations of Physics, 29*(1), 91-132. doi:10.1023/a:1018869221019

Metz, C. (1974). *Film language: A semiotics of the cinema*: University of Chicago Press.

Mitchell, D. (1983). The tuning of Harry Partch's 43 tones-to-the-octave just intonation scale, and it's musical endowments and consequences. *The journal of the Acoustical Society of America, 74*(S1), S22-S22.

mmorpg. (2016). Eve Online. Retrieved from http://www.mmorpg.com/gamelist.cfm/game/14/EVE-Online

mobygames. (2016). Eve Online. Retrieved from http://www.mobygames.com/game/eve-online

Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: from motion to'e'motion. *Social cognitive and affective neuroscience, 1*(3), 235-241.

Moscovich, V. (1997). French spectral music: An introduction. *Tempo*, 21-27.

Müller, M. (2015). Interactive Ambient Music Generation. In A. Berndt (Ed.), *Works in Audio and Music Technology*: TUDpress.

Neumann, K. (Writer). (1950). Rocketship X-M: Lippert Pictures.

Nierhaus, G. (2009). *Algorithmic composition: paradigms of automated music generation*: Springer Science & Business Media.

Nolan, C. (Writer). (2014). Interstellar: Paramount Pictures, Warner Bros.

Nordahl, R., & Nilsson, N. C. (2014). The Sound of Being There *Oxford Handbook of Interactive Audio*: Oxford University Press.

O'Callaghan, C. (2014). Auditory Perception. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* <http://plato.stanford.edu/archives/sum2014/entries/perception-auditory/%3E. (Summer 2014 Edition ed.).

Ólafsson, K. CALMUS. Retrieved from http://erkitonlist.is/kjartan-olafsson/

Olivier Deriviere, J. G., James Hannigan, Christian Henson, Jesper Kyd, Joris de Man, Garry Schyman, Joe Thwaites. (2013-15) *Unpublished* *interviews/Interviewer: J. Broomhall*. Game Music Connect 2013, 2014 and 2015.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The timbre toolbox: Extracting audio descriptors from musical signals. *The journal of the Acoustical Society of America, 130*(5), 2902-2916.

Penderecki, K. (1960). Anaklasis. Hermann Moeck Verlag - Celle.

Penderecki, K. (1961). Threnody to the victims of Hiroshima : for 52 strings = Threnos : den Opfern von Hiroschima : für 52 Saiteninstrumente. Melville, N.Y. : Belwin-Mills.

Penderecki, K. (1962). Flourescences. Hermann Moeck Verlag - Celle.

Persichetti, V. (1961). *Twentieth-century harmony: creative aspects and practice*: WW Norton & Company Incorporated.

Phillips, W. (2014). *A composer's guide to game music*: MIT Press.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *The journal of the Acoustical Society of America, 38*(4), 548-560.

Pochmursky, C. (Writer). (2009). My Musical Brain: Matter of Fact Media.

Polák, J. (Writer). (1963). Ikárie XB-1 Voyage to the End of the Universe: Filmové studio Barrandov.

Potty, G. R., & Parameswaran, V. (2015). *Investigation of acoustical parameters in the South Indian musical intervals.* Paper presented at the Proceedings of Meetings on Acoustics.

Price, H. (2013). Time's Arrow and Eddington's Challenge. In B. Duplantier (Ed.), *Time: Poincaré Seminar 2010* (pp. 187-215). Basel: Springer Basel.

Pöppel, E. (1978). Time perception *Perception* (pp. 713-729): Springer.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in cognitive sciences, 1*(2), 56-61.

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1525), 1887-1896.

Risset, J.-C. (1991). Timbre et synthèse des sons. *Le Timbre, métaphore pour la composition*, 239-260.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research, 3*(2), 131-141.

Roads, C. (2004). *Microsound*: MIT press.

Schaeffer, P. (1966). Traite ́des objets musicaux. *Paris: Editions du Seuil*.

Schafer, R. M. (1977/1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*: Destiny Books.

Schwarz, D., Beller, G., Verbrugghe, B., & Britton, S. (2006). *Real-time corpus-based concatenative synthesis with catart.* Paper presented at the 9th International Conference on Digital Audio Effects (DAFx).

Schwinger, W. (1989). *Krzysztof Penderecki: his life and work: encounters, biography and musical commentary*: European Amer Music Dist Corp.

Scorsese, M. (Writer). (2010). Shutter Island: Paramount Pictures.

Scott, R. (Writer). (1979). Alien: Brandywine Productions, Twentieth Century-Fox Productions.

Slater, M. (2003). A note on presence terminology. *Presence connect.* Retrieved from http://publicationslist.org/data/melslater/ref-201/a note on presence terminology.pdf

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3549-3557.

Smalley, D. (1997). Spectromorphology: explaining sound-shapes. *Organised sound, 2*(02), 107-126.

Spielberg, S. (Writer). (1977). Close Encounters of the Third Kind: Columbia Pictures Corporation.

Stannard, R. (2013). Second Law of Thermodynamics. In A. L. C. Runehov & L. Oviedo (Eds.), *Encyclopedia of Sciences and Religions* (pp. 2106-2106). Dordrecht: Springer Netherlands.

Steinitz, R. (2003). *György Ligeti, Music of the Imagination*: Faber and Faber Limited.

Stern, D. N. (2004). *The Present Moment in Psychotherapy and Everyday Life (Norton Series on Interpersonal Neurobiology)*: WW Norton & Company.

Szelag, E., Kanabus, M., Kolodziejczyk, I., Kowalska, J., & Szuchnik, J. (2004). Individual differences in temporal information processing in humans. *Acta neurobiologiae experimentalis, 64*(3), 349-366.

Tarkovsky, A. (Writer). (1972). Solaris: Creative Unit of Writers & Cinema Workers, Mosfilm.

TheFreeDictionary. (2016). The Free Dictionary Retrieved from http://www.thefreedictionary.com/envelope. Retrieved September 1st http://www.thefreedictionary.com/envelope

Truax, B. (1982). Timbral Construction in" Arras" as a Stochastic Process. *Computer Music Journal, 6*(3), 72-77.

Truax, B. (2001). *Acoustic communication* (Vol. 1): Greenwood Publishing Group.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306.

Vaggione, H. (1984). The making of Octuor. *Computer Music Journal, 8*(2), 48-54.

Valve. (2007). Portal: Valve Corporation.

Wackermann, J. (2007). Inner and outer horizons of time experience. *The Spanish journal of psychology, 10*(1), 20.

Waterworth, J., & Waterworth, E. (2013). Distributed Embodiment: Real Presence in Virtual Bodies.

Weir, P. (Writer). (1993). Fearless: Spring Creek Productions, Warner Bros.

Wierzbicki, J. (2014). The Imagined Sounds of Outer Space. *Journal of Sonic Studies*(8). Retrieved from Studies,https://www.researchcatalogue.net/view/109621/109622/0/0

Wilcox, F. M. (Writer). (1956). Forbidden Planet: Metro-Goldwyn-Mayer.

Wilder, B. (Writer). (1945). The Lost Weekend Paramount Pictures.

Wittmann, M. (2011). Moments in time. *Frontiers in integrative neuroscience, 5*, 66. Retrieved from http://dx.doi.org/10.3389/fnint.2011.00066

Wooller, R., Brown, A. R., Miranda, E., Diederich, J., & Berry, R. (2005). A framework for comparison of process in algorithmic music systems.

Xenakis, I. (1953/1954). Metastaseis B. Boosey & Hawkes.

Xenakis, I. (1955/1956). Pithoprakta. Boosey & Hawkes.

Xenakis, I. (1958). Concret PH (pp. Electroacoustic composition on tape).

Xenakis, I. (1992). *Formalized music: thought and mathematics in composition*: Pendragon Press.

Zemeckis, R. (Writer). (1997). Contact: Warner Bros.

Aarseth, E. (2008). A Hollow World: World of Warcraftas Spatial Practice. *Digital culture, play, and identity: A World of Warcraft reader*, 111.

APPENDICES

Appendix A. Prototype - NowEngine

The prototype algorithm, NowEngine, can be found in the Multimedia folder on the DVD that accompanies the thesis. The multimedia material and prototypes included on the DVD are meant for demonstration purposes only. They are not published along side the thesis. Future versions may be, however, and I therefore kindly ask readers not to pass the prototypes on to others without permission.

I have programmed the Max MSP[52] prototype, NowEngine, with the purpose of ”freezing” music or other sound in a *non-cyclically static* sustained *vertical expression*. It demonstrates well some concepts developed in the thesis including *micro-texture, cyclic* and *non-cyclic stasis, vertical expression, vertical extension, striated* and *smooth space-time* as well as issues concerned with the *complexity of the sounding musical manifestation*. Additionally, NowEngine may be setup to function as a simple 16-voice *VE-transformational* algorithm for the computer generation of *irreversibly changing* VDM.

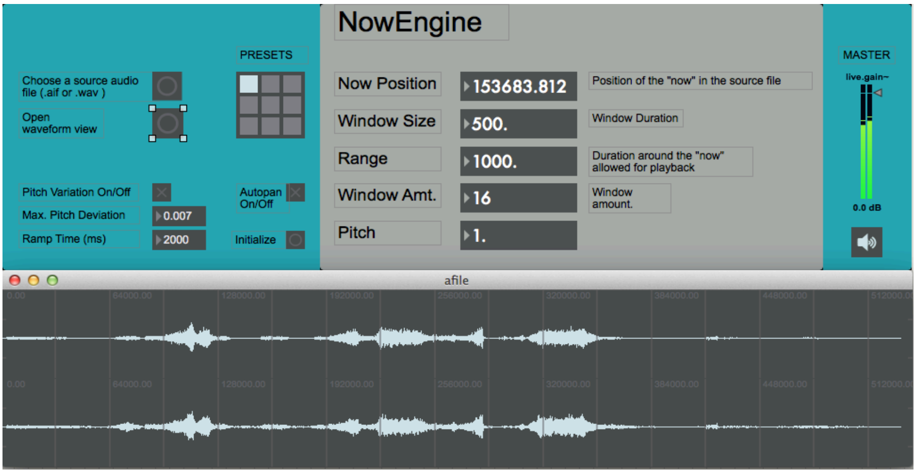


Figure: NowEngine GUI

It uses a technique inspired by granular synthesis in which envelope ”windows” fade the amplitude of a source audio file up and down in a overlaying manner. What I here refer to as ”windows” are the NowEngine’s equivalent to the ”grains” of granular synthesis. The main difference between the concept of grains and the envelopes created in NowEngine is the availability of longer durations of the

---

52 NowEngine was programmed using the modular programming environment Max7 from the company Cycling74.

envelope windows. A ”freeze” effect here maintains the micro-textural qualities of the source audio file - something which is better achieved when the generated envelope windows are allowed to exceed the micro time scale and enter the vertical extension time scale (i.e. achieved i NowEngine by setting the [Window Size] parameter to a value above 150 ms.) Another feature of NowEngine, which is not standard in granular synthesis, is its ongoing random positioning of new windows within the source audio file. This feature is particular important because it allows for what I have called *non-cyclic stasis* (see chapter 2) - a percept that is fundamentally different from *cyclic stasis*. The method for experimenting with cyclic and non-cyclic stasis is explained in the section about the [Range] parameter below.

The movements of the amplitude envelopes (windows) resemble the movements of a piston engine where the number of cylinders represent the amount of running envelope windows. For each window instance a random position within the [Range] of the [Now Position] is chosen in the source audio file. Playback of the instance now starts at this position while the amplitude is faded in and out according to the [Window Size]. The procedure starts over when the enveloped window's duration has past. Increasing the [Window Amount] to, for instance, **4** will activate a number of 4 window instances within the duration set by the [Window Size] parameter. The more windows the smoother becomes the vertical expression.

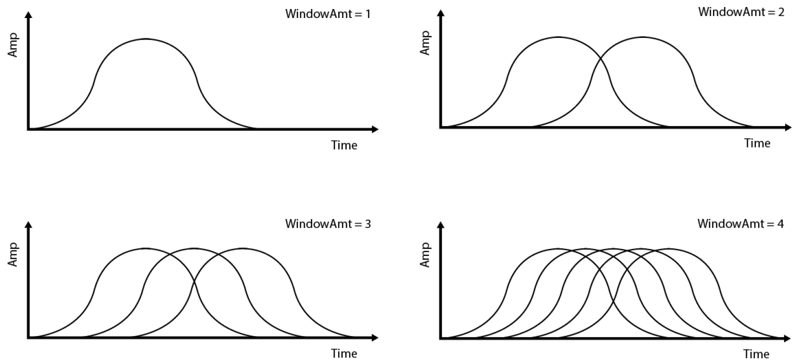


Figure: The envelope curves of 1, 2, 3 and 4 windows.

A runtime[53] version of NowEngine can be found in the multimedia folder on the DVD handed in with the thesis.

[53] i.e. a version that cannot be opened and further programmed. I do not include the accessible Max MSP patch as the inclusion of the prototype in the thesis is meant for demonstration purposes and should not be regarded a publication of the system as such.

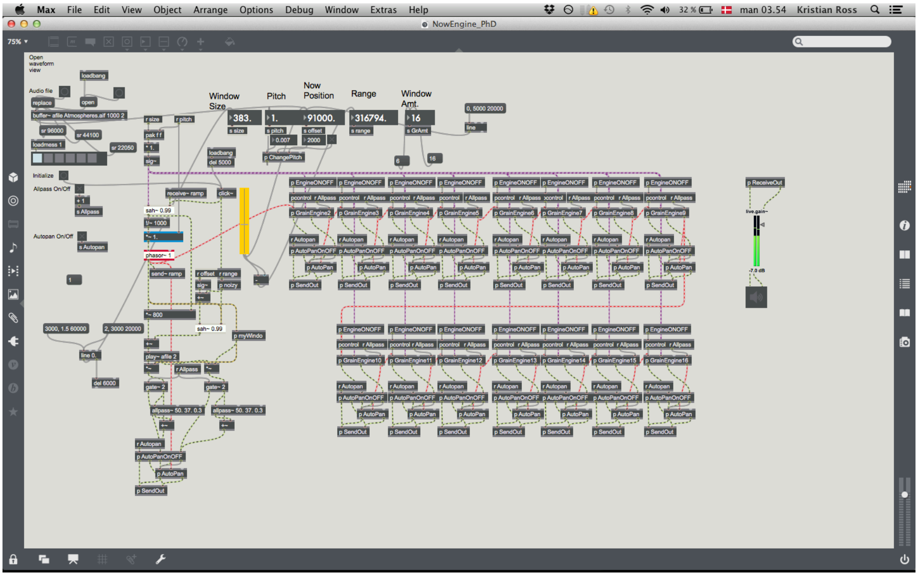


Figure: Excerpt from the Max MSP patch inside NowEngine.

The primary variable parameters of the system involved with the shaping of the vertical expression are:

- [Now Position]
- [Window Size]
- [Range]
- [WindowAmt]
- [Pitch]

The system also features functionality for pitch variations through the following parameters:

- [Pitch variation On/Off]
- [Maximum Pitch Deviation]
- [Ramp time]

And an autopan function:

- [Autopan On/Off]

All parameters can be set by click-and-dragging the number boxes as well as by clicking followed by entering a value manually. The [Now Position] can also be set by opening the [waveform view] and clicking anywhere in the source audio file waveform.

Now Position

The [Now Position] refers to the specified center position in the source audio file around which playback positions for each window instance are chosen. The position of each window instance is controlled indeterministically by the [Range] parameter. A Now Position of **0** represents the beginning of the source audio file. Higher values move the center position of playback - the *now* - further into the source audio file. This parameter is measured in milliseconds.

Window Size

The [Window Size] parameter sets the duration of the windows. This duration has a big influence on the degree of micro-textural information allowed to exist within the generated vertical expression. Durations within the micro time scale (shorter then approximately 150 ms) effectively produce what in the discourse of granular synthesis is referred to as ”grains” of sound. Longer durations ensure more elaborate micro-textures, but may also make audible any abrupt changes in the source audio file. This may cause separate gestalts to form and challenge the integrity of the vertical expression.

(This parameter of the prototype, when changed, causes the pitch to be rather erratic for a moment. This an unwanted artifact that quickly subsides.)

Range

The [Range] parameter sets the duration in milliseconds around this center positions at which windows may be played back. At a range of **2000 ms**, for example, windows are positioned at random within a range of 1000 ms (1 second) on each side of the [Now Position]. The range parameter is useful to test the perceptual implications of *cyclic* and *non-cyclic* stasis. Any setting **higher than 0** creates a non-cyclic static expression while a setting of **0** causes all windows to play back at the same specified [Now Position] - causing audible cyclic repetition. The perceptual difference between the two is quite clear. Test it out with a [Window Size] setting of, for example, 300 ms and a [Window Amount] of 10 with the [Pitch Variation] turned OFF.

Window Amount (WindowAmt)

This parameter sets the number of window instances activated within the duration of 1 Window Size. The system can operate with an amount of 1-16 window instances. A setting of **1** will cause an audible pulsating effect. A setting of **2** will double the frequency of the pulse. The windows are overlapping, so as the [Window Amount] increases the smoother becomes the pulsating as the individual enveloped sound events perceptually merge into the greater whole - a sound-mass. At settings of **4** and higher the pulsating effect perceptually disappears and a smooth-spaced vertical expression remains (i.e. when the [Range] parameter is above **0** so as to ensure non-cyclic stasis.)

Pitch

With the [pitch] parameter you can change the pitch at which the source audio file is played back. This works across all window instances.

Pitch variations

The [Pitch variation On/Off] toggle switch activates an animation of the [Pitch] parameter to gradually change to a randomly chosen value position within the range set by the [Maximum Pitch Deviation] parameter. Each new [Pitch] value is reached within the duration set by the [Ramp time] parameter.

Autopan

The system operates in stereo and the stereo image of the source audio file is per default maintained in the generated vertical expression. However, NowEngine features an autopan option by which each new window instances may be placed at random positions in the stereo field. This can result in a ”cloud-like” effect as the windowed audio events are ”sprinkled” across the stereo image.

Start-up procedure

To get started with generating vertical expressions follow this start-up procedure:

- Switch DSP on by clicking the [Speaker Icon] at the bottom of the MASTER section to the right in the NowEngine user interface.
- Open a source audio file of your choice with the [Choose a source audio file] button to the left in the NowEngine user interface.
- Open the waveform view by clicking the [Open waveform view] button. Click somewhere in the waveform.
- [Initialize] - if nothing works, press this button.

Presets

Presets can be selected by clicking the preset matrix at the left in the user interface. Settings are stored as a preset by holding [*Shift*] while clicking a preset slot.

- **[Preset 1]** is an effective starting point for the generation of vertical expression. Open, for example, Ligeti’s *”Atmosphêres”* (found in the NowEngine folder on the DVD) and press [Preset 1]. Open the waveform view by clicking the [Open waveform view] and click around in the waveform of *”Atmosphêres”*. Changing the [Now Position], [Window Size] and [Range] parameters will allow for a wide variety of vertical expressions to by generated from *”Atmosphêres”* - preserving the micro-textural properties.
- **[Preset 2]** is set up to as a VE-transformational algorithm. Open *Str_5secLegato-5secTremolo_D6-B-1.wav* one the NowEngine folder to hear this. See below.

Irregular Orchestra sound library

As is covered in Appendix B, I used the NowEngine to produce the samples of a sound library for introducing irregularities on a per sample basis in my collaboration CCP and Kjartan Ólafsson on EVE Online and CALMUS. This was done using the included *Str_5secLegato-5secTremolo_D6-B-1.wav* sound file along with similar sound files for brass and woodwind instruments[54]. These sound files feature all the pitches of their instrument group's register range chromatically moving down from the highest possible tone to the lowest possible tone. Additionally, these pitches are represented in two different textures. The mentioned string group sound file, for example, consists of a sequence that starts with the pitch **D6** played sostenuto for 5 sec, then with tremolo for 5 seconds - it then moves a semitone down for 5 seconds sostenuto and 5 seconds tremolo and so on. This continues for all available pitches in the instruments group ending, in the case of the string instrument group, with the low **B-1**.

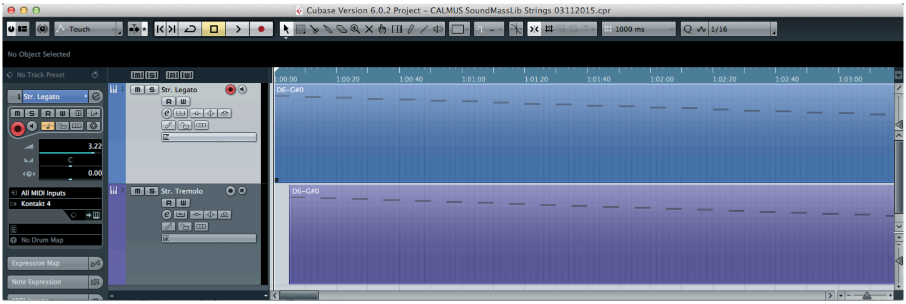


Figure: Showing part of the Cubase arrangement used to produce the Str_5secLegato-5secTremolo_D6-B-1.wav sound file.

The sound library was produced by setting NowEngine to hold a [Now Position] in the source audio sound file corresponding to the center position of all the pitches one by one and recording the output of each resulting tone. With a [Range] of **10000 ms** (thus including both the sostenuto and the tremolo part as possible start positions for the randomly chosen window instances) and with active [pitch variation] this produced a series of samples that are irregular in terms of texture and pitch. Additionally, by setting the [WindowAmount] to around **7** also subtle amplitude variations start to become perceivable. This procedure was repeated for all three instrument groups until all pithes of the groups were represented in the sound library.

[54] I produced these sound file in Steinberg Cubase by mixing third party sound libraries I had access to.

NowEngine as a VE-transformational algorithm

NowEngine can function as a simple VE-transformational algorithm with adaptive potential. Here every new window instance is seen as a vertical expression. The generated vertical expressions are in turn combined into an ongoing irreversibly changing VDM sound-mass. In order to hear what this sounds like with a string orchestra, click the [Choose a source audio file] and in the NowEngine folder of the multimedia folder choose the file called: *Str_5secLegato-5secTremolo_D6-B-1.wav*. Then press [Preset 2] in the ”Preset” box.

Try to play around with especially the [Range] and [Now Position] parameters to adaptively change the characteristics of the generated sound-mass. Here the [Now Position] acts as a register parameter (higher values generate lower pitches), while the [Range] controls the register range.

The NowEngine prototype may be expanded with more textures, instruments and other features. Its basic functionality presented here, however, has proved to be a very good starting point for VE-transformational VDM experimentation as well as for various sound design purposes. Apart from the work I have done with EVE Online and CALMUS, the presentation of these endeavors are beyond the focus of the thesis.

Appendix B. EVE Online, CALMUS and the "IrregularOrchestra" Sound Library

Introduction

In this appendix, I present a practical case study concerned with the realization of VDM through the algorithmic composition program, CALMUS with the aim of implementation in the Icelandic space-based MMO, EVE Online. The appendix describes the process and results of experiments done in the context of my ongoing research collaboration with lead audio designer at CCP, Baldur Baldursson, and inventor of CALMUS, Professor Kjartan Ólafsson.

This collaboration is focused around an aim to implement CALMUS in EVE Online as a means for algorithmic music composition. However, as is made evident throughout the thesis, generative music in games has a number of aesthetic challenges, some of which may be regarded as particularly significant to VDM (such as ensuring a sufficient *complexity of the sounding musical manifestation*, highlighted in chapter 8). My involvement is primarily concerned with addressing these issues by making possible the generation of aesthetically viable VDM in the game through CALMUS. A part of this involvement thus entails the investigation of potential methods for adapting CALMUS to this purpose but bears an equal potential for also enriching the sounding complexity of any other music composed by this system.

The problems and criteria of success highlighted here represent my own experiences and ideas, which have been supported, inspired and adapted through the many discussions I have had with Kjartan Ólafsson and Baldur Baldursson. They are not backed up by an academic literature review, but originate from addressing practical issues. Additionally, the efforts I have made in the collaboration to solve a range of challenges by experimentation and practical prototyping has inspired some of the realizations and ideas in the thesis (such as realizing the importance of the sounding manifestation, seeing first hand the aesthetic challenges of generative music as well as understanding its great potential for adaptive game music).

The part of our collaboration described in this appendix does not deal with an actual implementation of music in EVE Online and issues such as trans-diegetic couplings of triggers, cues and RPCs to musical parameters are not discussed. (Such experimentation work is are briefly covered in Appendix E.) Rather, I present here a number of identified challenges connected to the general implementation of CALMUS in EVE Online with a focus on solving identified aesthetic issues. The aim is to allow the output of CALMUS' generative algorithm to reach an artistic quality sufficient for implementation in high profile AAA games such as EVE

Online. That being said, the results of the experiments are highly applicable outside the specific context of EVE Online and CALMUS.

The first part of the appendix serves as a short introduction to EVE Online's myth and gameplay as well as to the basic principles and functionalities of CALMUS. The wish to implement CALMUS in EVE Online has demanded solutions to a number of concrete problems associated with the use of samplers to manifest the generated music. I have identified these problems to be primarily concerned with issues of *regularity* on the sound object, vertical extension and micro time scales. These problems are highlighted in this appendix and possible solutions are presented. During my collaboration with CCP and Kjartan Ólafsson, I proposed that the problems of regularity may be approached by two directions: a *per sample* approach and a *sound engine* approach. The first in concerned with building a sample library with irregular sounding audio files. The second introduces irregularity through the process of playing back sound samples. My experiments led to the production of an irregular sound library. They also led to the conclusion that a sound engine-based manifestation is better suited in terms of adaptability and capability to solve more of the presented regularity issues.

A three-part construction was proposed, where CALMUS composes structural data in the form of MIDI messages or OSC data, a sound library provides the digital sound material with which these musical structures are rendered and a custom sampler allows for the superimposing of irregularity on the vertical extension and micro time scales. Many of my research results have led to a planned adaptation of CALMUS to include features proposed in this appendix. These new features are listed at the end of the appendix.

EVE Online

The following section is a short description of some significant narrative and gameplay related aspects of EVE Online.

EVE Online is a massively multiplayer online role-playing game (MMORPG) initially released in 2003 by Icelandic game development company, CCP. The game takes place in a science fiction based game world about 20.000 years into the future. The player takes on the role as a spaceship pilot pursuing fame and success in a vast game world consisting of over 5000 star systems. Here the player is presented with an immensely multifaceted gameplay in the worlds largest coherent game world in which all players are playing within the same universe on the same servers. The game features a dynamic player-controlled economy and the gameplay is characterized by the so-called *sandbox* concept. Sandbox games have no predetermined storyline for the player to move through and the game primarily offers only content and functionality on the basis of which the players alone are responsible for the cause of events.

At the time when I first got involved with EVE Online and CCP, the game exhibited only linear music in the form of a jukebox in which players could also play music of their own choice. At the time of writing, however, adaptive music of an ambient character has been implemented to a limited degree in a part of the game world called *wormhole space* via the functionality offered by Wwise. Initial steps have been made for the implementation of CALMUS in the game although it has not yet at this early stage made it into a released game expansion.

For a short video presentation of CALMUS composing music inside a test platform[55] of EVE Online, see the video ”CALMUS EVE ONLINE ex2” in the multimedia folder on the DVD. The content of this video has been produced without my involvement, but gives an impression of the game and a hint of CALMUS' capabilities as well as a sense of some of the regularity issues described later in this appendix.

Figure: In-game image from EVE Online.

The EVE Online myth and gameplay

The game's myth describes how a natural wormhole was discovered, which led to a previously unknown part of the universe. Ports were established so that explorers could travel back and forth between the old and the new world, which was named Eve. The portal suddenly and violently closed, leaving behind a number of settlers cut off from returning to the Milky Way civilization. Of those who survived the harsh conditions in the new world, five unique tribes arose. After generations of war the player enters the new world of Eve in a period of relative peace between empires

[55] A parallel copy of the game used for development purposes but not accessible by the consumer player.

and is free to choose an identity within one of the tribes as well as pick one of approximately 50 different careers.

In the document, EVE Careers Guide (CCP) a number of possible careers are mentioned. Each of these careers represents different types of gameplay although many gameplay elements are also shared between careers. It mentions 49 careers organized into 5 overall categories. A large number of different corporations have been formed by the players during the years EVE Online has been playable online and a large part of the game's plot has become to fight over resources, make diplomatic deals with other corporations and tribes, mining asteroid belts for resources, trade, plundering other space pilots for resources, blowing up spaceships for fun and, not least, fleeing for your life in the huge game world. A significant aspect of the game is the butterfly effect of player actions caused by the fact that all players are situated within the exact same game world, on the same servers.

All of the different gameplay types in EVE Online - however different - entail an alternation between taking place either inside space stations or outside in outer space. Another important aspect of the game across careers is the fact that different star systems in the enormous game world have different so-called *security statuses.* This is measured for each system by a scale of 11 subdivisions spanning from 0.0 to 1.0. A security status of 1.0 means that the star system is practically completely safe to travel in, while a security status of 0.0 on the other hand is lawless territory in which the risk of being met by hostile opponents is very high.

One of the first tests I took part in was concerned with the implementing of VDM through Wwise. Here these game states were chosen for the initial trans-diegetic coupling to the adaptive music. This experiment is described briefly in Appendix E.

CALMUS

CALMUS (or CALMUS) is an algorithmic compositional tool developed by Kjartan Ólafsson. The program makes possible the formalization of contemporary compositional methods through artificial intelligence based on cellular automata[56]. In this section I will present some of CALMUS' basic functionality. I argue that CALMUS is capable of composing VDM with its existing functionality but also suggest areas in which CALMUS might be expanded in order to better accommodate a more radical approach to composing VDM.

The following is a short description of some of the most significant compositional functions that can be carried out by CALMUS as well as a brief overview of the hierarchy and basic concept of the program. The macro-structure, the larger musical form, of a piece being generated by CALMUS is organized manually by placing so-called *musical objects* on a sequencer track.

[56] Cellular automata is described in chapter 8 and allows for the continuous modification of cells and offers the possibility for real-time manipulation of the generated music.

A musical object contains melodic material constructed by so-called *melodic cells* and harmonic material based on harmonic controlling profiles or prototypes called *model chords*. A melodic cell is composed by the organization of *tones* either manually or generatively by CALMUS. These tones are defined by four basic components; *pitch, attack time point, duration* and *velocity*.

These components lay the foundation for the basic musical representation of the program - a typical trait of generative algorithms as was highlighted in chapter 8. This representation positions CALMUS (in its current version) amongst many other generative systems by having a basic unit of composition that resides within the sound object time scale.

Melodic and harmonic material within musical objects is arranged musically by one of four available polyphonic modes; *individual voices, homophony, horizontal-vertical voices* and *canon voices*. The mode, individual voices, generates independent melodic material for each voice. These independent voices can, however, relate harmonically. In the mode of, homophonic voices, every voice is given the same attack time position and duration, whereby the voices are temporally synchronized. Horizontal-vertical voices offers the opportunity to distribute tones of a melodic line over several voices - a method often used in twelve-tone music where it is often referred to as horizontal-vertical disposition. The last of the four polyphonic modes is canon voices. Here the voices share certain characteristics but are divided by variable durations constructing a canon-like musical polyphonic structure in which the user has control over parameters such as *time regularity, frequency of disposition* and *ratio of truth*. The possibilities of canon structure within CALMUS is described in considerable detail in the CALMUS tutorial (Ólafsson).

Other basic parameters in CALMUS include choice of 2nd, 3rd, 4th, 5th, 6th and 7th harmony, 31 different musical scales, as well as functionality for orchestration based on traditional and temporary orchestral schemes for the symphonic orchestra’s main instrument groups: Strings, woodwinds, percussion and brass instruments.

The musical material generated by CALMUS is output as data in a form, ideal for conversion into MIDI messages or used through the OSC protocol. This data is comprised of the four mentioned components of a *tone* in CALMUS: pitch, attack time point, duration and velocity. As is discussed in chapter 8 generative algorithms have the sound object time scale as basis for its basic unit of composition, the musical note. CALMUS is no different in this respect in its current state.

Notes on CALMUS in use

It is clear that CALMUS is capable of generating textural music with its current functionality. Structural schemes such as micro-polyphony and formation and development of tone clusters and other meso sound-mass approaches are all within the scope of the program. The program’s orchestration capabilities and polyphonic modes allow vertical expressiveness within the confinement of pitch-space and possibilities for irregularity within the meso time scale.

CALMUS is capable of utilizing both gameplay-dependent and gameplay-independent mappings. At this stage it offers adaptive methods for horizontal development by allowing data inputs to control the sequential order of its musical objects, instrumentation and orchestration changes as well as gradual transitions between different composition scheme presets (it is e.g. possible to smoothly change from a homophonic preset to a canon preset). The program's internal sequencer additionally allows for deterministic gameplay-independent horizontal development by allowing musical objects to be manually positioned on a timeline - each object comprising its own orchestration paradigm, instrumentation, polyphonic mode, harmony, register and so on. Although both the gameplay-dependent and gameplay-independent methods are available at this point in CALMUS, it depends primarily on the implementation of the program's functionality in a game context how much control is handed to the gameplay.

However, in order to offer better control within the vertical expression time scale as well as within spectral space additional functionality is needed. CALMUS features no functionality at this point that allows access to the vertical extension and micro time scales or which can lift the striation of the vertical dimension offering access to the micro tonal world of spectral space.

Problems and solutions

CALMUS is a *generative* algorithm and outputs a score-based musical representation typical for generative systems. It is based on cellular automata enabling it to offer realtime manipulation of compositional parameters. As is discussed in chapter 8. This makes it suitable for use in games because the character of the musical output can be adaptively altered on the fly according to the gameplay. However, the system's aesthetic capabilities in regards to the *complexity of the sounding musical manifestation* as well as in regards to *musical coherency*[57] also follow the general trend of generative systems. This means that the quality of the sonorous expression tends to be insufficient if CALMUS is set to play directly through a sampler with no other measures taken in terms of music-shaping stages.

During my involvement with the project it was agreed that initial experiments should be carried out based on a *sampler* approach where the manifestation of the musical representation is performed by a sampler program rather than through synthesis. The reason for this is that CALMUS offers extensive orchestration functionalities that are designed based on an acoustic orchestral paradigm. In order to take advantage of this functionality a sampler loaded with orchestral sounds was deemed most relevant for testing.

---

[57] See chapter 8.

However, a sample-based manifestation of a score-based musical representation poses sonorous inadequacies of its own that are closely tied to certain issues of *regularity* in the manifestation process. A key aspect of enriching the complexity of the sounding musical manifestation is thus to introduce *irregularity* into the manifestation of the musical representation. During the project I proposed two approaches to meeting this challenge: The introduction of irregularity on a:

- *per sample basis* by designing a library of sound files that feature irregularity on the vertical extension and micro time scales
- *sound engine basis* by manipulating the way samples are played back

Both of these approaches are discussed in the following pages. First, however, I must deal with the concrete problems of irregularity posed by a sample-based manifestation method.

Implications of a sample-based manifestation

Modern digital sample libraries let today's composers and sound designers access the timbral and expressive powers of real-world sounds such as that of the symphonic orchestra. These libraries come in many versions and qualities and sometimes include a sound engine (a sampler program) for sample playback with differing functionality and expressive diversity and at varying price tags. A very popular software engine for sample playback and manipulation is Native Instruments Kontakt (NI Kontakt) in which third party developers can build custom ”instruments”, which include the functions they wish to associate with their sound library (such as a custom designed parameter for changing the pedal noise in NI Kontakt's *Upright Piano* patch or the virtual positioning of the microphones used to record the samples of a drum set in Toontrack's *EZ Drummer* Kontakt Instrument). Sounds that were previously unaffordable to many composers because of the pricy endeavor of hiring an orchestra of real musicians are by now common inventory in even the most basic home studios.

As part of an in-the-box[58] composition workflow discussed in the section on music-shaping stages in chapter 8, these libraries often exhibit a certain homogeneous or ”clinical” sonic feel, which is mainly caused by *regularities* in the perfectly pitched, recorded and produced samples. This perfection causes many libraries to sound ”unnatural” compared to the lively, infinitely nuanced and highly *irregular* quality of a real orchestra of musicians playing together. The typical regularity problems of sound libraries I address here exist at the vertical extension and micro time scales as well as within the sound object of the musical tone. Examples of problematic regularity issues that can occur in the practical use of sample libraries include:

1. Pitch is perfect
2. Pitch is consistent

[58] I.e. entirely produced within the confines of a computer.

3. Amplitude envelope is consistent across samples
4. Spectral evolution is similar across across samples
5. Articulation and playing techniques are ON/OFF - not gradual
6. Key-mapping may use same samples across several pitches decreasing sonorous complexity
7. Note repetition uses same sample or a limited *round robin* approach
8. Long attack times work well with notes of long duration, but poorly with short notes.
9. Fast music sounds worse than slow music

Problem 1

That the pitch intonation is perfect becomes a problem of regularity especially in polyphonic situations where the spectral flux created in the polyphonic harmony is lessened by the very precise intonation.

Possible solution: The issues of perfect pitch can be solved both by creating a sample library where samples vary slightly in pitch in an irregular way as well as by allowing the pitch to be continuously altered by the sound engine that plays the samples.

Problem 2

Typically pitch intonation is consistent during a tone and either conforms to the equal-tempered tuning or other tuning schemes. This causes similar issues of regularity. When pitch variation is introduced by the sampler playing the samples, the available functionality for pitch deviation is often *regular* - such as LFO's[59] - although random generators exist in some samplers. **Possible solution**: This issue may likewise be solved by both approaches.

Problem 3

Also the amplitude of a tone may be consistent in the sustain of the tone or may follow a static amplitude envelope. Attack times of note onsets can thus become perceivably identical. This occurs horizontally with sequences of notes. It can also happen vertically when all notes of a particular instrument patch are played with the same amplitude envelope setting.

Possible solution: A per sample approach can feature amplitude variations during the duration of the sample, but cannot solve the issue of identical amplitude envelopes. Here, a random choice of envelope within a definable range could be implemented through the sound engine.

Problem 4

---

[59] Low Frequency Oscillator, which produces one of several possible wave forms at a frequency typically between 2 and 20Hz.

The spectral evolution within tones (i.e. the timbral evolution of the spectral envelope for the duration of the sound object) is overly simplified compared to live musicians as articulations of tones are homogenous across pitches of the same sample patch.

Possible solution: If the samples of a sound library have a constantly varying spectral evolution this could solve the homogeneity across pitches, but when the same note is repeated, a per sample approach falls short. Here a sound engine approach would allow for the same note to not have an un-identical spectral evolution every time it is played.

Problem 5

Playing techniques and articulations (such as tremolo, legato or sforzando) are either on or off and gradual expression change is typically not easily accomplished.

Possible solution: A per sample solution to this entails implementing in the same sample an alternation over time between different playing techniques. This works in terms of introducing irregularity, but does not offer adaptive control over when for instance tremolos take gradually over from a legato articulation. While this may be irrelevant in the name of VDM's *un-specificity* it ignores the potential for coupling articulation and playing technique adaptively to game events. A sound engine approach could allow such couplings.

Problem 6

A very typical source for regularity is associated with key-mapping - that is, the mapping of samples across the virtual keyboard. The same samples are often mapped within to cover more than one semitone (or microtone). This is done to keep the library size down, but the result is that the very closely related versions of the same sample (only differing by being played at a slightly different speed and therefor pitch) lowers the perceived complexity of the expression and introduces yet another layer of regularity between tones.

Possible solution: In a per sample approach this represents a need to have at least as many samples as there are notes in the tone system. It seems more difficult to solve this particular problem through a sound engine-based approach to introducing irregularity.

Problem 7

Note repetition is often confined to the same sample - unless so-called round-robin is available as a function by which a selection of sample variations (additional samples) help to avoid sample repetition. This rarely includes alternating between more than two samples. When a note is repeated in a voice and triggers the exact same sample it becomes obvious to the ear that the sonic expression is harnessed within a limited frame set by the involved technology. In other words, the strongly limited expressiveness of the machine as compared to a real orchestra reveals itself through obvious repetition.

Possible solution: This is not possible to solve on a per sample basis. A way of addressing this problem through a sound engine, however, could be to randomly change the onset position of sample playback at every trigger. In order to avoid phase issues, this should be done either by picking one of a number of positions which are sufficiently far apart or by having a rule associated with the random generator that states not to have positions too close to each other play back simultaneously.

Problem 8

Another typical issue, when playing musical structures through samples is that long attack settings in the dynamic envelope work well with notes of long duration and vice versa. When trying to play back shorter notes, a long attack will counteract the articulation of the note.

Possible solution: A possible work-around to this would be to introduce duration-dependent attack time of amplitude envelope so that short durations result in short attack times and vice versa. This is possible only through a sound engine approach to introducing irregularity.

Problem 9

The problems of computer-generated music, especially when pursuing an acoustic sound ideal, become more obvious the more attacks and releases of tones the music presents. In faster passages in which the density of tone onsets and terminations is relatively high, these aesthetic shortcomings, which are not least caused by a lack of variation and irregularity, are therefore more significant to the musical expression than in slower passages.

Possible solution: As it is to a large extent exactly the attack and release part of the dynamic envelope of tones that are problematic, VDM, in which these aspects of the sonorous music is hidden, obscures such shortcomings featuring sometimes primarily the continuation part of the tones. Synthesized sounds sometimes work better in the limited expressive conditions of computer-generated music because they are generally speaking more ”aligned” with the mechanical aesthetics of the limited score-based musical representation.

Technical implications of sample-based manifestation

The size of a sound library in terms of disk space depends on the length and resolution of the samples, the resolution of key-to-sample mapping in regards to both pitch and dynamics[60], and the amount of diversity of instrumental playing techniques. As disk space and CPU power allocated to sound are generally scarce commodities in games, any solution to address the problem of regularity must take

[60] i.e. how many audio samples are available to occupy the keyboard as well as the different velocities at which each key may be played.

these aspects into account. A *per sample* implementation of irregularity requires a relatively long sample length to keep the listener from recognizing repetitions as the sample loops. This increases the necessary amount of disk space.

If irregularity is introduced instead as part of the sound engine, sample length can be kept much shorter and disk space requirements can be kept low. This approach could on the other hand result in an increased CPU load depending on the method chosen - for example when many samples are played at the same time, are subject to heavy DSP effects or where a granular technique is used such as in the prototype, NowEngine (Appendix A).

The ”IrregularOrchestra” sound library

Sample-based irregularity is relatively easy to implement without having to change CALMUS’ generative algorithm. As an initial test I therefore produced an irregular sound library. Irregularity was implemented in the audio files on the parameters of pitch, amplitude and micro-texture (articulation). This was done using the prototype, NowEngine. The sound library features three instrument groups: strings, woodwinds and brass instruments. It has only a single sample per pitch to represent all velocities. However, all semitones have their own dedicated sample. The samples are 20 seconds long to avoid a sense of repetition in the irregular micro-variations as the sample loops. As mentioned each sample varies in terms of pitch, amplitude and micro-texture. In regards to micro-texture string samples vary irregularly between *sostenuto* and *tremolo*, the brass samples vary between the pure tone of *piano* playing and the distinct ”overdrive” sound of *forte,* and the woodwinds alternate irregularly between *vibrato* and *non-vibrato*.

To read about how the sample library (titled ”IrregularOrchestra”) was produced see Appendix A. To listen to the samples see the ”B - IrregularOrchestra” folder on the DVD accompanying the written thesis. I subsequently used the ”IrregularOrchestra” sample library to compose four VDM pieces, ”DistanceMusic 1-4” (also present in the multimedia folder), which can be consulted for an auditory and practical insight in many of the principles of the thesis. The ”DistanceMusic” pieces are described in Appendix C.

These VDM pieces where in turn cut into smaller segments and layers and used as test material for the prototype, KPS (KeynotePlaybackSystem), which functions as a transformational algorithm through random reorganization of the ”DistanceMusic” fragments. KPS is exemplifies quite well the flexibility and adaptability of VDM because although not taking any measures to fit the ”DistanceMusic” fragments together in any particular way (other than being split into a high and a low register) the vertically dominated musical output seems to function aesthetically well at all times. It is advised to leave the KPS system running for half an hour or so to experience an irreversibly changing flow of VDM. Some repetition may become perceivable after some time due to the relatively low number of fragments used.

Nevertheless, the system serves the purpose as a proof of concept prototype acceptably.

Solutions on a per sample and a sound engine basis

The irregular sample library was relatively easy to produce and quickly accomplished a degree of irregularity. However, a sample-based approach has a number of limitations. First of all, its irregularities cannot be controlled adaptively as it does not allow access to either the sound object, vertical extension or micro time scales. Additionally, the samples have to be relatively long, which increases the amount of disk space that has to be allocated to sound in the game. This can become a rather big issue when more instruments, articulations, velocities and playing techniques are added.

The introduction of irregularity on a based on the sound engine that plays back the samples would on the other hand be able solve all of the above problems and at the same time allow access to adaptively control the amount and type of irregularity in real-time - eventually by gameplay-dependent mapping to triggers and RPCs in EVE Online. The main challenge here is that a sound engine-based approach requires the generative algorithm to output data that can represent such micro-variations and sound object time scale envelopes. This surpasses the current score-based representation capabilities of CALMUS and entails a potentially large programming task.

Nonetheless, a three-part construction was proposed as the most promising long term solution. Here CALMUS would compose musical data represented as MIDI messages or OSC[61] data, a sound library would provide the digital sound material with which this musical representation is rendered and a custom sampler would allow for the superimposing of irregularity on the sound object, vertical extension and micro time scales. Additionally, the proposal importantly includes the need for an adaptation of CALMUS to expand its representational capabilities by allowing both static and adaptive envelopes to be output for the control of the custom built sampler's irregularity functionality.

Concluding remarks

A wish to implement CALMUS in EVE Online using sampling as the sound generating method has prompted my investigation into some general issues of a sample-based musical manifestation. These issues stem mostly from a lack of irregularity at the sound object, vertical extension and micro time scales. I have proposed two approaches to addressing these challenges, a *per sample* approach, in which irregularity is implemented in the audio files of a sample library, and a *sound*

[61] Open Sound Control (OSC) is a protocol similar to MIDI but optimized for modern networking technology.

engine approach, in which irregularity is implemented during the process of playing back the audio samples.

The outcome of my involvement and investigations concluded that within a sampler-based manifestation paradigm a sound engine approach for introducing irregularity to the sounding musical manifestation is the most fruitful endeavor of the two mentioned approaches. This conclusion is based on balancing CPU and disk space consumption with aesthetic and adaptive gain. My practical experiments and reflections have led to the current development of CALMUS to include an initial expansion of its musical representation with adaptive envelopes of pitch and amplitude micro-variations as well as for manipulations on the sound object time scale. These new aspects of the representation are meant to control a custom built sampler that includes the functionality proposed as solutions to the problems of regularity mentioned in this appendix. The functionality planned for this initial adaptation of CALMUS include:

- Subtle pitch and amplitude fluctuations (on the micro and vertical extension time scales) for each voice or every MIDI channel, which can be indeterministically or stochastically mapped to in-game data streams (i.e. random within a specifiable range or based on probability calculations).
- Static and adaptive envelopes (on the sound object time scale) for at least each MIDI channel, ideally all voices, that works on a *per note* basis and can control amplitude, pitch or any other accessible vertical parameter.
- Adaptive data for the start position of samples in the custom built sampler.
- Duration-dependent attack time of amplitude envelope on a per note basis

Appendix C. DistanceMusic Production

This appendix describes four linear VDM pieces titled ”DistanceMusic 1-4”, which can be found in the multimedia folder accompanying the thesis.

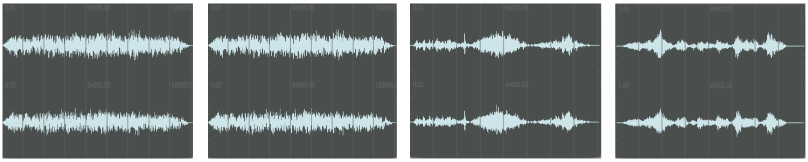


Figure: DistanceMusic 1, 2, 3 and 4

Initial purpose

I originally composed the pieces to be experimentally implemented in a mysterious region called ”Wormhole Space” in the immense EVE Online game world. Here their purpose was to play whenever the player zoomed out completely while flying around in a spaceship - giving him or her an overview over the situation, and in turn removing the player visually from the immediate action close to the space ship. Due to certain insufficiencies in the middleware product Wwise in terms of VDM implementation, and due to the fact that CCP did not have the resources allocated at the time to program a specialized VDM based system, the implementation of the DistanceMusic was put on hold together with other VDM experiments in Wwise (presented briefly in Appendix E). However, the pieces represent an aesthetic that we collaboratively assessed was suitable for EVE Online’s mysterious Wormhole Space - an aesthetic closely tied to the structural characteristics of VDM.

Production notes

The four DistanceMusic tracks were produced in Steinberg Cubase mainly using the IrregularOrchestra sound library discussed in Appendices A and B. Using this sound library ensured a degree of irregularity and a richness to the complexity of the sounding manifestation that surpassed for example the NI Kontakt 4 Factory Library. However, quite a lot of additional automations were made manually to pitch and amplitude as well as several other vertical parameters such as filtering, various DSP effects, articulation and instrumentation.

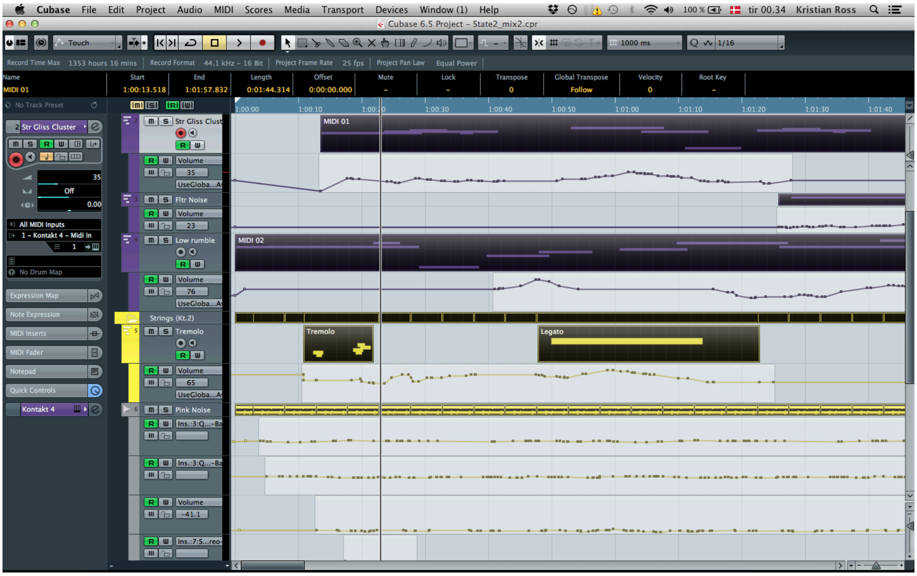


Figure: Showing part of the Cubase arrangement window for "DistanceMusic" production.

Musical characteristics

DistanceMusic 1-3 are highly dissonant and are characterized by aggressive articulations, dark timbres and a general tendency for low register content and some degree of micro-tonal harmonies. DistanceMusic 4 stands out from the other pieces by being the only representative of what I have referred to in chapter 6 as *benign VDM*. Its harmonies are more consonant and the piece positions itself further towards historically based anticipation (or past-listening) than the others - without, importantly, ever becoming horizontally dominated.

All four DistanceMusic pieces exhibit an orchestral instrumentation and exhibit the characteristics[62] of VDM by: Avoiding the formation of horizontally separable gestalts through obscuring the onset and termination of tones as well as focus on the mass-merging of many individual events; avoiding perceivable horizontal regularities through the absence of repetition, rhythmic pulse and accentuation of meter as well as by using patterns like themes, motifs and rhythmic patterns only very scarcely; and lastly, by weakening of the bond between spectromorphologies by avoiding to a large extent causal direction and by exceeding the period of event cohesion. However, they contain occasional horizontally dominated traits and may for instance give rise to a subtle sense of melody, harmonic progression and causal direction. This is especially true for "DistanceMusic 4".

[62] i.e. these are, I must recapitulate, *characteristics* and are not to be understood as categorically followed rules of VDM.

A key aspect of the compositional process has been to implement irregularities manually, and it is the experiment's perhaps largest realization that in order to make a sampler-based manifestation of VDM sound natural, irregularities on the micro, vertical extension, sound object, and meso time scales are an absolute necessity.

DistanceMusic fragments as source material for KPS

As mentioned the DistanceMusic pieces are all linear. Furthermore, they have been composed manually - by a human PhD student. During my work on the transformational algorithm KPS (KeynotePlaybackSystem) I needed material that the system could transformationally reorganize into an irreversibly changing flow of VDM. I therefore cut the DistanceMusic pieces into 31 stems and fragments of 30-60 seconds for KPS to randomly reorganize both horizontally and vertically. The result became a good demonstration of VDM's *musical flexibility* and *adaptability* as the output of KPS, which was based on 31 fragments of all DistanceMusic pieces, seemes to function musically well regardless of which fragments were layered or succeeding each other. The KPS system is described in Appendix D.

Appendix D. KeynotePlaybackSystem (KPS)

The prototype program, KPS, can be found in the Multimedia folder on the DVD that accompanies the thesis. The multimedia material and prototypes included on the DVD are meant for demonstration purposes only. They are not published along side the thesis. Future versions may be, however, and I therefore kindly ask readers not to pass the prototypes on to others without permission.

KPS is an experimental transformational algorithm for adaptive VDM. It is loaded with a selection of musical material comprising meso-structural fragments cut out of the linear pieces ”Distance Music 1, 2, 3 and 4” presented in Appendix C. A runtime version of KSP can be found in the multimedia folder of the DVD accompanying the thesis. I developed the KPS prototype based on experiences from working with Wwise. This prototype, in fact, represents most of the functionality that we (i.e. The sound department and CCP and myself) were missing in Wwise in order to be able to implement adaptive irreversibly changing VDM in EVE Online.

The relatively small number of fragments used means that some sense of repetition will occur while listening to the system. This is is especially true for the more horizontally dominated passages, which are, clearly, more memorable. This problem can be solved by including a much larger reservoir of vertically dominated musical meso-structures into the prototype.

The adaptive functionality of the prototype currently allows for control over mainly two states (High and Low Threat level) as well as variables associated with its four indeterministic envelope generators. However, the prototype could be expanded in the future to compose in a VE-transformational manner through duplications of the already existing functionality and through the loading of a larger number of vertical expressions.

KPS works in two states: ”Low Threat” and ”High Threat”. (These are meant to represent two possible game states in EVE Online, which are associated with certain regions in the game that have different threat levels. See Appendix B for a brief overview of EVE Online’s gameplay and myth.)

KPS as demonstration of theory

This prototype demonstrates well the *musical flexibility* of VDM. As mentioned, the musical material that is transformationally recombined by KPS comprises fragments from the linear DistanceMusic pieces. Apart from having organized these fragments into two register categories (high and low) their vertical and horizontal recombination happens completely arbitrarily - and the vertically dominated musical manifestation does not suffer under this. This ”careless throwing together” of musical fragments demonstrates equally well the *unspecificity* of VDM discussed in chapters 4 and 8.

The fragments from DistanceMusic pieces contain horizontally dominated traits and these are audible from time to time in the flow of VDM output from KPS. These horizontally dominated traits theoretically keeps the listener in a state of musical listening, which is further promoted by the instrumentation (symphonic strings) being highly associable to music.

The two states, high and low threat, as well as all of the parameters of the envelope generators, represent possible narrative couplings to game play and games states as covered in chapter 6. Additionally, the indeterministic manner in which the envelopes function may counteract such trans-diegetic coupling becoming *overt*.

KPS also demonstrates the use of *gameplay-independent envelopes* for horizontal development of VDM as described in chapter 9.

Practical use

To get it running fast, wait for a minute after opening, then press [Start] in the upper right corner.

The GUI of KPS is divided into 4 main sections (brown colors) as well as a MASTER section at the right (dark grey color). The two left-most brown sections control the amplitude envelope generators for the ”Low Threat” state. The upper ”Low Threat” section controls ”high register” musical material, the lower section controls ”low register” musical material. This scheme is identical in the two right-most brown sections, only here the state is ”High Threat”. KPS can be set adaptively to run in either the ”High Threat” or the ”Low Threat” state. This can be done manually or automatically.

Figure: The GUI of KPS

This prototype features a number of variables that can be utilized adaptively in a game context. At this time, however, they are changeable mainly by hand for demonstration purposes. Below is an overview of the available parameters and their functions.

Variable Parameters

MASTER section

- [Start] - starts the prototype.
- [Stop] - stops KPS at the end of the next fragment.
- [Change state when next fragment is triggered] - pushing this (a big red button with yellow circle) will make KPS change its threat status to either ”Low Threat” or ”High Threat”, depending on which state it is already in. The transition happens with the next triggered sample, so it might take a few seconds before any change becomes audible.
- [Auto State Change On/Off] - turns on and off automatic state change.
- [Automatic State Change Interval] - these controls allow KPS to change automatically between threat levels at a regular time interval.
- [Pitch shift +/- Semitones] - sets the the pitch range at which fragments are played back.
- [Speaker Icon] - Start and stop audio processing here.

Envelope Generators

(Change the ranges by click-and-dragging the yellow and red fields)

- Attack Range - sets the range in which a random attack time is chosen for the amplitude envelope of fragments
- Hold Range - sets the range in which a random hold time is chosen for the amplitude envelope of fragments
- Release Range - sets the range in which a random release time is chosen for the amplitude envelope of fragments
- Pause Range - sets the range in which a random pause time is chosen between each new amplitude envelope is triggered
- Envelop meter (Env.) - here you can monitor the movement of the amplitude envelope

The envelopes do not have complete control over the amplitude of the output VDM as the fragments themselves contain amplitude variations. Additionally, there is a short fade silence in between each fragment. For more control, the audio material used by KPS should have a more consistent amplitude.

Meters

The large meters to the left in each section of the prototype (Labeled ”Input Gain”) show the current amplitude of the fragment playing. This is the only way, at the moment, to monitor visually which threat level is active. When the meters at the left are active, it is the Low Threat level that is playing. The High Threat level is associated with the ”Input Gain” meters in the two sections to the right in the prototype.

Presets

- Presets can be selected by clicking the preset matrix at the left in the user interface. They are stored by holding [*Shift*] while clicking a preset slot. The default is preset 1.

NB: The envelope generators behave erratically for the first minute after opening the prototype. This is an unwanted artifact that I will correct sometime in the future. (The KPS prototype runs on a mixture of ectoplasm and dark matter - both of which have to warm up for a minute or so. I am looking into solutions to the issue.) Wait for about a minute and press [Start] in the top right corner of the prototype.

Appendix E. Wwise and Wormhole Space

In 2014 I was involved in an experiment with implementing VDM in a test version[63] of EVE Online through Wwise. The music was intended for a special region of the game world called ”Wormhole Space” (W-Space). Wwise is a widely used middleware program for implementing adaptive sound and music in games and it is capable of doing this task well in most cases. As we soon found out, however, it is not very well suited for VDM.

In this appendix I very briefly provide an overview of some key aspects in the experiment and identify insufficiencies in Wwise in regard to adaptive VDM implementation.

Wormhole Space

The region in EVE Online called ”Wormhole Space” represents a number mysterious regions that players can travel to by flying their spaceships through wormholes. W-Space differs from the rest of the game world. The laws of physics are different in W-Space, making the spaceships’ internal systems behave strangely in regard to, for example, ”weapon recharge duration” and ”defense shield power”. This strange effect on the spaceships is determined by the *class* of wormhole space in conjunction with the *type of wormhole* that was used to get there. See below. The mysterious qualities of this part of the game world seemed (and *seems* on the basis of the research of this thesis), very well suited for VDM with its unique semiotic and perceptual attributes.

Different classes of W-Space exist in EVE Online, some are relatively safe to travel to and others are dangerous. W-space is inhabited by *sleepers* - a special and very dangerous class of NPCs (Non-Player Characters). Sleepers are unmanned drones that belong to a mysterious old race. So far only drones have been observed, but their presence suggests the existence of a civilization that has not yet been seen. The higher the W-Space class, the tougher and more dangerous are the sleepers.

Trans-diegetic coupling concept

The experiment entailed implementing a background *VDM keynote* (i.e. a subtle, irreversibly changing, vertically dominated background music) on which *incidentals* (short VDM meso-structures) could be triggered when a player would accidentally awaken a sleeper - and be in grave trouble. Since the class of W-Space determines the danger level, keynotes of *benign* VDM would provide a sense of ”awe and wonder” in the less hazardous classes while more sinister VDM would give a sense of ”danger and tension” in dangerous W-Space classes. (Four of the ”Sleeper

[63] A parallel EVE Online game that is used for testing at CCP and is not accessible to players.

awakens" incidentals that I composed for the experiment can be found in the multimedia folder accompanying the thesis.)

Wwise VDM insufficiencies:

It quickly became clear that Wwise, while being highly adept for implementing adaptive HDM, did not have the necessary functionality for adaptive VDM. The primary insufficiencies were associated with a lack of *indeterministic*, *gameplay-independent* implementation functionality. While it is possible too use Wwise to sequentially organize musical segments at random, it does so on the basis of musical *meter.* It is not possible to deactivate musical meter as the main clock to which any change must adhere. Furthermore, changes had to be made *specifically* - that is, any change to the musical flow had to be specifically programmed either to *gameplay-dependent* RPCs or as *static envelopes* that were tied to the meter grid. The program, due its clear focus on HDM, did not offer unspecific, indeterministic, gameplay-independent functionality. The functionality we were missing in Wwise in order to successfully implement constantly varying VDM keynotes can be summed up to the following:

- Indeterministic envelopes for irregularity and variation
- Functionality for unspecific musical development
- Gameplay-independent meta-structures for irreversible change

The experiences from the experiment led to the realization that the above functionality are, while not necessarily difficult to program, not available in Wwise due to its focus on HDM. This prompted the development of the prototype KPS, which is presented in Appendix D. KPS (KeynotePlaybackSystem) features precisely the functionalities, whose absence prevents Wwise from being suitable for adaptive VDM.

Appendix F. Vertical Parameters

This appendix provides a selection of vertical parameters that were mentioned in the thesis and which form part of the structural basis for an initial set of compositional principles for adaptive VDM in computer games. The parameters are divided into five categories: Tone systems, harmonic parameters, timbral parameters, textural parameters and other parameters.

Tone systems

- Pythagorean tone system
- Pure tuning
- Tempered tone systems such as meantone temperament, 12 tone equal temperament and other equal tempered octave divisions
- Micro-tonalities (quarter tone, eights tones etc.)
- Open spectral continuum (partial-based orchestral synthesis)

Harmonic parameters

- Pitch
- Scales and modes within the chromatic 12 tone equal tempered striation paradigm
- Musical key (Polytonality, bi-tonality and anchoring.)
- Harmonic bass interval: 2nd, 3rd, 4th harmony etc.
- Cluster harmony
- Register
- Register range
- Consonance/dissonance

Timbral parameters

- Spectral envelope
- Spectral centroid
- Spectral spread
- Spectral deviation
- Noisiness
- Harmonicity / inharmonicity
- Spectral stretching and compression
- Spectral flux
- Instrumentation and orchestration (herein acoustic instruments as well as other sound generating techniques such as various forms of synthesis and sampling.)

Textural parameters

Articulation and instrumental expression:

- Tremolo, vibrato, flutter tongue etc.
- Penderecki's expansions notated as: *molto vibrato*, *very slow vibrato with 1/4 tone frequency difference*, *very rapid non-rhythmical tremolo*, *repeat tone as rapidly as possible*, *repeat the notated tone groupings as rapidly as possible* as well as *to saw* and *to rub*.
- DSP effects
- Pitch modulation
- Amplitude modulation

Textural motion (Smalley):

- Continuous - discontinuous texture motion
- Iterative <-> granular <-> sustained
- Periodic <-> aperiodic
- Acceleration, deceleration and flux
- Saturate and granular noise
- Cyclic auditory stasis
- Non-cyclic auditory stasis

Other parameters

- Occupation of spectral space by: Emptiness-plenitude, diffuseness-concentration, streams-interstices, and overlap-crossover
- Number of auditory streams (governed by gestalt principles of *auditory stream segregation*)
- Properties of *aural space:* Reverb type, room size, pre-day, decay time, and difussion.

Appendix G. Contents of the multimedia DVD

Appendix A - NowEngine

- NowEngine prototype
- Atmosphêres.aif
- Str_5secLegato-5secTremolo_D6-B-1.wav

Appendix B - EVE Online and CALMUS

- CALMUS EVE ONLINE ex2.mp4
- ”IrregularOrchestra” sound library samples

Appendix C - DistanceMusic

- DistanceMusic1.wav
- DistanceMusic2.wav
- DistanceMusic3.wav
- DistanceMusic4.wav

Appendix D - KPS

- KPS prototype
- Audio folder (containing fragments from the four DistanceMusic pieces)

Appendix E - Wwise and Wormhole Space

- SleeperAwakens_3.wav
- SleeperAwakens_6_2.wav
- SleeperAwakens_7.wav
- SleeperAwakens_9.wav

VDM Pioneers

- Ligeti - Atmospheres.aif
- Ligeti - Lux Aeterna.mp3
- Penderecki - Threnody for the victims of hiroshima.mp3
- Xenakis - Metastaseis.mp3

Appendix H. Glossary

The glossary presented in this appendix contains a list of proposed additions to the terminology of the research field. These terms and concepts have been developed through the thesis out of necessity. In some cases additions were prompted by the unique implications of addressing VDM. In other cases they are suggested as a means for clarifying existing terminology. The glossary can thus be seen as a list of contributions to the research field, each pointing to its own theoretical context elaborated in the thesis.

- *Arborescent structure* - a branching, tree-like structuring scheme (borrowed from Deleuze and Guattari.)
- *Auditory horizontality* - a perceptual phenomenon of those auditory developments that exceed the ”vertical extension” and which facilitate the formation of clear temporally truncated gestalts.
- *Auditory verticality* - a perceptual phenomenon, which is facilitated by an illusion of ”auditory stasis” caused by structural properties of a ”vertical expression” and occurring within a temporal window of a ”vertical extension.”
- *Complexity of the sounding musical manifestation* - the complexity in terms of phrasing, articulation, acoustics and vertical expression associated with the sounding musical expression as heard by the listener.
- *Cyclic auditory stasis* - an auditory phenomenon associated to cyclic repetitive sound organizations within the vertical extension time scale, constituted by the continuous repetition of micro-textural regularities.
- *Cyclic musical stasis* - a music-perceptual phenomenon associated with cyclic repetitive sound organizations at the meso time scale, where the continuous repetition of musical substructures conveys a sense of status quo.
- *Direct adaptive sound and music* - refers to adaptability that is based on direct player input rather than, for instance, indirect game states. The term is synonymous to Karen Collins’ notion of *interactive dynamic sound and music.*
- *Empathetic co-agency* - a defining factor of *musical listening,* as I conceive it. A music perceptual phenomenon that refers to the linguistic, motor-action, musical and emotional representation of the *intentions* of a *sender* as experienced by a receiver.
- *Equilibrium-based anticipation* - an idealistic anticipatory listening position in which the listener is very familiar with the music and which involves a stable cognitive equilibrium and is encouraged by musical structures that exhibit low musical entropy, causal direction, horizontal dominance, much repetition, and in which events take place through cyclic repetition.

- *Future-listening* - see *equilibrium-based anticipation*.
- *Game-external competences* - refers to media competences that are established outside the context of the specific game that is being played.
- *Game-internal competences* - refers to media competences that are established inside the context of the specific game that is being played.
- *Gameplay-dependent development* - horizontal musical development that is facilitated by data streams coming from insight the game.
- *Gameplay-independent development* - horizontal musical development that is facilitated by mechanisms that are independent of data streams coming from insight the game.
- *Horizontally dominated music (HDM)* - music in which the expressivity of *auditory horizontality* dominates over that of *auditory verticality*.
- *Historically based anticipation* - an omnipresent anticipatory listening position in which the listener is unfamiliar or partly familiar with the music. It involves an alternation between cognitive perturbation and equilibrium and is encouraged by musical structures that exhibit repetition, belong to a recognizable style, contain causal direction and musical structures that are recognizable or easily learned.
- *Indirectly adaptive sound and music* - refers to adaptability that is based on data streams coming from the game such as RPCs rather than on direct player inputs. The term is synonymous to Karen Collins' notion of *adaptive dynamic sound and music.*
- *Irreversible change* - musical metaphor for the quality of continuous temporal forward movement whose goal does not let itself be predicted because the qualities of the musical structure hint towards a future with an infinite number of possible states.
- *Irreversibly changing vertical musical time* - an appendix to Kramer's *vertical time*, denoting a perceptual *non-linear* musical time conception in which temporal development clearly plays a part in the musical expression, but does so with no predictable goal direction. Encompasses the sense of time associable to, for instance, Xenakis' *"Metastaseis"* (1953-54) and *"Atmosphêres"* (1961).
- *Micro-generative algorithm* - a music compositional algorithm that generates *vertical expressions* and composes music based on the the *micro time scale* as its lowest basic unit of composition.
- *Music* - from a perceptual perspective: *that,* which is experienced through a state of *musical listening*, and which involves *empathetic co-agency*.
- *Music-shaping stages* - refers to the differing shaping processes that a *sounding musical manifestation* has gone through, and in which differing degrees of complexity is added to the expression.

- *Musical entropy* - refers to disorder in terms of harmony, melody, rhythm and timbre.
- *Musical flexibility* - the capability of music to be sequentially reorganized while maintaining musical compatibility between adjacent sections. *Rhizomatic* musical structures have a high musical flexibility.
- *Musical listening* - an expression of *empathetic co-agency* as defined in the thesis.
- *Non-cyclic auditory stasis* - an auditory phenomenon associated to non-cyclic sound organizations within the vertical extension time scale, constituted by micro-textural irregularities.
- *Non-cyclic musical stasis* - associated with non-cyclic sound organizations at the meso time scale that are not perceptually reducible into constituents of horizontal regularities.
- *Now-listening* - see *perturbation-based anticipation.*
- *Past-listening* - see *historically based anticipation.*
- *Perturbation-based anticipation* - an idealistic anticipatory listening position in which the listener is unfamiliar with the music. It involves constant cognitive perturbation and is encouraged by musical structures that exhibit high musical entropy, no causal direction, vertical dominance, no repetition, and in which events take place through irreversible change.
- *Rhizomatic structure* - a non-linear structuring scheme in which all point connect to all other points (borrowed from Deleuze and Guattari.)
- *Undisclosed sender* - associated with the undisclosed ”origin” of linguistic, motor-action, musical and emotional *intentions* as these are co-enacted by the listener through *musical listening*.
- *Vertically dominated music (VDM)* - music in which the expressivity of *auditory verticality* dominates over that of *auditory horizontality*.
- *VE-transformational algorithm* - a music compositional algorithm that recombines *vertical expressions* and has the *vertical extension time scale* as its lowest basic unit of composition.
- *Vertical expression* - an emergent hybrid concept of pitch, harmony, timbre and micro-texture, which occupies the *vertical extension*.
- *Vertical extension* - the subjective and relative duration within which a listener experiences auditory verticality.

SUMMARY

The PhD research investigates the implications of implementing adaptive VDM (Vertically Dominated Music) in computer games. VDM is proposed as a term for music that is perceptually dominated by vertically oriented (simultaneously sounding) musical qualities like harmony, timbre and micro-texture and encompasses musical styles such as sound mass music, textural music and spectral music.

The perceptual and music-psychological implications of introducing VDM in games and VDM's narrative potential in a multi-modal context are examined. Here the focus is on VDM's association with narrative phenomena such as outer space, infinity, mystery, divinity and paranormal activity by making sensory accessible the presence of "something" that exceeds the boundaries of the senses.

The thesis argues that a specialized system for adaptive VDM demands a fundamentally new approach to computer-generated adaptive game music and suggests a structural, phenomenological and functional framework to form the theoretical basis for the development of such a system.

ISSN (online): 2246-123X
ISBN (online): 978-87-7112-821-5

AALBORG **UNIVERSITETSFORLAG**